NEURAL MECHANISMS AND DYNAMICS UNDERLYING REACHING AND DECISION MAKING

A DISSERTATION
SUBMITTED TO THE PROGRAM IN NEUROSCIENCE AND THE COMMITTEE ON GRADUATE STUDIES OF STANFORD UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Matthew T Kaufman
July 2011
I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Krishna Shenoy, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

John Huguenard

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Eric Knudsen

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Tirin Moore

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.
Abstract

The ultimate purpose of the motor system is clear: it exists to control the body. However, despite the motor system being among the longest-studied brain structures, it remains unclear how – mechanistically – motor cortex performs this function. Here, a mechanistic approach was taken to investigate how primary motor cortex (M1) and dorsal premotor cortex (PMd) control movement. That is, the goal was to elucidate the dynamics of the motor cortex ‘machine.’ Monkeys were trained in reaching tasks, and neural signals were recorded from their brains as they performed them. Two broad classes of analysis were used. First, cell-by-cell analyses were combined with cell-type analyses, which permitted examining the activity patterns of excitatory and inhibitory neurons separately. Second, techniques based on dynamical systems analysis (such as dimensionality reduction) were applied, which permitted analysis of neural populations as a whole and abstraction to a somewhat higher level of system function. Three major results and a technical advance are presented. Firstly, we investigated how it is possible for an animal to hold still even as neural activity in motor cortex changes drastically during preparation for the upcoming movement. We found that, contrary to common assumptions, there does not appear to be a ‘gate’ comprised of high inhibition during preparation. Instead, using the dynamical systems perspective, we found that preparatory activity has a special structure such that it remains in intrinsically muscle neutral, ‘iso-force’ patterns. Secondly, we searched for coherent dynamics in the movement-time activity of motor cortex. We found that motor cortex appears to obey a relatively simple set of dynamics, dominated by oscillatory patterns. Moreover, the exact neural trajectory is heavily determined (‘seeded’) by the immediately preceding preparatory activity. In order to causally perturb these
dynamics with patterned stimulation and cell-type specificity, we then developed a
set of optogenetic techniques for use in primates. Finally, we investigated how the
dynamics of the decision-making process are reflected in motor cortex. To do so,
we combined a novel decision-making paradigm, many simultaneous neural record-
ings, and single-trial analytical techniques. Preliminary results are given for this final
section, demonstrating the presence of vacillation in monkey decision-making. In sum-
mary, we found that preparation and movement can be understood as an oscillatory
dynamical system seeded by preparatory activity that lives in an iso-force space, that
inhibitory and excitatory neurons seem to play more similar roles in the dynamical
system than might be expected, and that moment-by-moment processes of motor
decision-making can be seen in motor cortex.
Acknowledgements

Good science is rarely performed in isolation, and I feel extremely lucky to have found the Shenoy lab. My advisor, Krishna Shenoy, has his students’ and post-docs’ interests at heart more than I could possibly have asked for, and has been generous with his time, thoughtfulness, career advice, and of course scientific thinking at every stage. I’ve felt from the beginning that I have a friend and advocate for an advisor. I was also incredibly lucky to end up working with Mark Churchland, with whom most of the experiments in this thesis were performed. He is a four-sigma scientist ($\pm 2\sigma$, $p < 0.001$) who has pushed me to attack the big problems and is relentlessly creative. He has contributed enormously to my scientific growth, usually intentionally. I am also thankful for Mackenzie Risch (our technician), who besides taking excellent care of our animals and helping to train me in their handling, did a lot toward making 4 years in a cramped basement doing experiments fun. I would also like to thank Ilka Diester, who was both careful and congenial in pursuing optogenetics; Stephen Ryu, who implanted my monkeys’ arrays; my committee members (John Huguenard, Eric Knudsen, and Tirin Moore), who have helped to focus me on the most important parts of my work; and the other members of the lab (past and present), who have become friends and have made the intellectual environment richer.

Finally, a few key personal acknowledgements. To my parents and grandparents, who have always encouraged me to do whatever it was I thought I would enjoy, always assumed I was capable of succeeding at it, and supported me in every way. And, of course, to Adriana, who has improved me as a person, has been my companion and intellectual partner, has kept me sane, and has made me happier than I thought I was constitutively capable of.
## Contents

Abstract

Acknowledgements

### 1 Introduction

1.1 Motivation ........................................... 1
1.2 Background ........................................ 2
1.3 Summary of thesis work ............................ 9

### 2 Lack of inhibitory gating in PMd

2.1 Introduction ........................................ 13
2.2 Methods .............................................. 16
  2.2.1 Subjects ........................................ 16
  2.2.2 Task apparatus ................................ 18
  2.2.3 Task design .................................... 18
  2.2.4 Neural recordings, classification of neuron types, and EMG .... 20
  2.2.5 Neural analyses ................................ 20
2.3 Results .............................................. 21
  2.3.1 Separation of narrow- and broad-spiking neurons ......... 21
  2.3.2 Comparison of the responsiveness of the two neuron classes .. 23
  2.3.3 Additional controls for neuron classification ............. 28
  2.3.4 Consistency of neural modulation across conditions between planning and movement .......................... 29
  2.3.5 Non-canonical waveforms ........................ 32
2.4 Discussion ........................................... 34
  2.4.1 Interneurons are more responsive than pyramidal neurons ... 34
  2.4.2 Lack of support for output gating by inhibition ............... 35
  2.4.3 Alternative mechanisms for preventing plan activity from driv-
           ing movement ........................................ 35
  2.4.4 Summary ........................................... 37

2.5 Supplementary Methods .................................. 38
  2.5.1 Neural recordings .................................... 38
  2.5.2 EMG recordings ..................................... 39
  2.5.3 Classification of neuron types using single-electrode recordings .... 40
  2.5.4 Classification of neuron types using multi-electrode array record-
           ings .................................................. 42

3 Control of functional connectivity in monkey motor cortex 44
  3.1 Introduction ......................................... 45
  3.2 Results .............................................. 47
    3.2.1 Behavior and recordings ............................. 47
    3.2.2 Separation of narrow- and broad-spiking neurons ............ 47
    3.2.3 Testing the feed-forward gating model .................... 50
    3.2.4 The iso-force space hypothesis of movement gating ......... 55
    3.2.5 Controls for the iso-force space analyses .................. 62
    3.2.6 Estimating the magnitude of the iso-force effect .......... 65
  3.3 Discussion ........................................... 65
    3.3.1 Evidence against the feed-forward gating hypothesis .......... 67
    3.3.2 The iso-force space hypothesis ........................ 67
    3.3.3 Summary ......................................... 69
  3.4 Methods ............................................. 70
    3.4.1 Subjects ......................................... 70
    3.4.2 Task apparatus .................................... 70
    3.4.3 Task design ...................................... 71
    3.4.4 Neural recordings, classification of neuron types, and EMG ... 71
4 Dynamics during preparation & movement

4.1 Summary ................................................................. 76
4.2 Introduction .......................................................... 77
4.3 Results ................................................................. 79
  4.3.1 Oscillatory responses in different motor systems .......... 79
  4.3.2 Firing rate oscillations are prevalent across neurons ...... 82
  4.3.3 Oscillations behave lawfully across conditions ......... 82
  4.3.4 Controls: jPCA isolates rotational structure only when such
      structure is present .................................................... 87
  4.3.5 Response heterogeneity is expected given underlying oscillations 92
4.4 Discussion ............................................................ 95
4.5 Methods .............................................................. 97
  4.5.1 jPCA ................................................................. 97

5 Optogenetics in monkey .................................................. 99

5.1 Introduction .......................................................... 100
5.2 Results .............................................................. 101
  5.2.1 Optogenetic inhibition ........................................ 103
  5.2.2 Optogenetic excitation ........................................ 105
  5.2.3 Effect of optical stimulation on passive movements ..... 108
  5.2.4 Bistable optogenetic excitation ............................. 109
  5.2.5 Histological evaluation of opsin expression ............. 110
5.3 Discussion ............................................................ 110
  5.3.1 Efficacy of optogenetics in non-human primates ....... 112
5.4 Methods .............................................................. 115
  5.4.1 Optical stimulations and neural recordings ............. 115
  5.4.2 Optical and electrical stimulation for movement analysis 116
  5.4.3 Data analysis ..................................................... 116
  5.4.4 Population PSTHs ............................................. 117
  5.4.5 Waveform analysis ............................................ 117
List of Tables

6.1 Behavioral choice probabilities ........................................ 123
6.2 Axis method decode accuracy at 100 ms after Go cue ................. 126
List of Figures

1.1 M1 and PMd locations .................................. 3
1.2 Reaction time vs. delay .................................. 4
1.3 Key situations for using single-trial analyses ............... 10

2.1 The output-gating hypothesis ............................. 15
2.2 Locations of narrow- and broad-spiking neuron recordings in PMd ....................................... 17
2.3 Delayed reach tasks ..................................... 19
2.4 Trough-to-peak histograms in PMd ........................ 22
2.5 Example PSTHs by cell type for PMd .................... 24
2.6 Comparison of response properties between narrow- and broad-spiking neurons in PMd ...................... 25
2.7 Cell-by-cell analysis of firing rates in PMd ............... 27
2.8 Relationship between plan and peri-movement cross-condition modulation patterns in PMd ................. 30
2.9 Analysis of preferred directions by cell type ............... 32
2.10 Analysis of ‘non-canonical waveform’ neurons ............ 33
2.11 Recording signal paths ................................. 39
2.12 Waveform alignment procedure ......................... 41
2.13 Trough-to-peak values by alignment method ............. 42

3.1 Behavior and example neuron and muscle responses ....... 48
3.2 Recording locations and neuron classification .............. 49
3.3 Feed-forward gating model and example recordings ....... 51
Chapter 1

Introduction

Motor cortex is involved in both the preparation and execution of movements. Yet, it has remained unclear how this preparatory neural activity relates to the subsequent neural activity during movement. My research has focused on elucidating this preparation-movement relationship, and on determining the key differences between these two types of neural activity. This thesis contains findings that preparatory activity is unlikely to be ‘gated’ by inhibition but avoids causing movement using special patterning; findings that the dynamics of activity during movement follow from the preparatory state and obey largely straightforward oscillatory dynamics; technical advances in bringing optogenetic tools to primates in order to perturb this activity; and findings that the moment-by-moment processes of decision-making are present in motor preparation, as evidenced by vacillation in the motor plan.

1.1 Motivation

The ultimate purpose of the motor system is clear: it exists to control the body. However, despite the mammalian motor system being among the longest-studied brain structures (Evarts, 1966, 1968; Fritsch and Hitzig, 1870; Leyton and Sherrington, 1917; Penfield and Boldrey, 1937), it remains unclear how – mechanistically – motor cortex performs this function. Fortunately, since motor cortical neurons are as few as two synapses from the muscles (Kuypers, 1960; Landgren et al., 1962), there has
been hope of relating the activity of these neurons to external variables such as muscle activity or movement kinematics.

A substantial literature exists examining the motor system using methods largely borrowed from studies of sensory systems (see Background, below). In traditional studies, monkeys have been trained to perform a movement task, individual neurons have been recorded during performance, then an attempt was made to relate the responses of the neurons to parameters of the movement (or, in some cases, the forces or joint torques involved). As discussed below, despite many careful studies, this approach has seen somewhat limited success.

Instead, it is possible to take more seriously the notion that the motor system exists to act as a controller for movement (Churchland et al., 2006b; Churchland and Shenoy, 2007b; Churchland et al., 2010a; Cisek, 2006b; Fetz, 1992; Scott, 2008; Todorov and Jordan, 2002). In doing so, a different class of question becomes primary. Instead of asking what parameters of movement neural responses relate to, one might wish to understand the operating principles of the controller—that is, what kind of dynamics it obeys, how the population of neurons functions together, the roles of different cell types, and how the brain computes what signal to produce. In other words, given that the patterns of muscle activity during movement are complex (Churchland and Shenoy, 2007b; Hatsopoulos et al., 2007; Scott, 2004, 2008), one might expect a complex control signal during movement, and ask how this complex signal can be generated. I have chosen to focus especially on how preparatory activity relates to the subsequent movement activity, in the hopes that this relationship would prove revealing of deep structure in the system.

1.2 Background

The motor system spans a number of brain areas across two lobes, frontal and parietal. My focus has been on two frontal areas, primary motor cortex (M1) and dorsal premotor cortex (PMd) (Figure 1.1). Our lab focuses on these areas because together they contain approximately 60% of the neurons that project down the spinal cord
Dum and Strick, 1991), and because they are generally thought to be the most critical areas for visually-controlled reaching (Boussaoud and Wise, 1993; Halsband and Passingham, 1985; Moll and Kuypers, 1977; Mushiake et al., 1991; Passingham, 1985; Tanji, 2001; Weinrich and Wise, 1982; Weinrich et al., 1984). Thus, they comprise the most direct brain outputs for controlling reaching.

In addition to being involved in controlling movement itself (Evarts, 1966; Leyton and Sherrington, 1917; Weinrich and Wise, 1982), M1 and PMd (among other areas) are also known to be involved in the preparation of movement (Crammond and Kalaska, 2000; Godschalk et al., 1985; Kurata, 1989; Messier and Kalaska, 2000; Riehle and Requin, 1989; Snyder et al., 1997; Tanji and Evarts, 1976; Weinrich et al., 1984). This process of movement preparation is not optional (e.g., Day et al., 1989; Ghez et al., 1997; Keele, 1968; Kutas and Donchin, 1974; Riehle and Requin, 1993; Rosenbaum, 1980; Wise, 1985). That is, when presented with a cue to move, we are unable to begin moving immediately. Instead, it appears that the brain must perform some time-consuming process before it sends commands to the muscles. If the upcoming movement is instructed during a delay period before the ‘Go’ cue, the brain can be given a ‘head start’ in performing these computations (e.g., Crammond and Kalaska, 2000; Riehle and Requin, 1989; Rosenbaum, 1980), thereby speeding the reaction time (RT; e.g., Churchland et al. 2006c; Figure 1.2). This preparatory process is thus often studied using a delayed-reach paradigm (see Figure 2.3).
Preparatory activity appears to play a vital role in producing the subsequent movement. Trial by trial, it predicts both RT (Bastian et al., 2003; Churchland et al., 2006c; Riehle and Requin, 1993) and variability in the subsequent movement (Churchland et al., 2006a), and its disruption delays movement onset (Churchland and Shenoy, 2007a). However, though it is known that preparatory activity is important and that it relates to the subsequent movement, the nature of this relationship has remained murky.

As noted above, a common assumption is that movement activity codes directly for some set of task or kinematic (joint) variables, and that preparatory activity simply constitutes a sub-threshold version of this activity (e.g., Bastian et al., 1998; Benjamin...
et al., 2010; Bullock and Grossberg, 1988; Cisek, 2006a). Under these assumptions, many studies have sought to identify the relevant variable or variables represented during preparation. Correlations have been found with a great many movement parameters, including direction (Riehle and Requin, 1989; Tanji and Evarts, 1976), distance (Messier and Kalaska, 2000; Riehle et al., 1994), reach speed (Churchland et al., 2006b), reach curvature (Hocherman and Wise, 1991), grip type (Goschalker et al., 1985), visual location of the target (Shen and Alexander, 1997), and eye position relative to the hand and target (Batista et al., 2007; Pesaran et al., 2006).

However, serious problems exist for this representational framework. First, the mere presence of tuning for these parameters is not enough to imply that they are what motor cortex ‘codes’ for; even muscles can have preferred directions (Mussa-Ivaldi, 1988). Second, tuning for these parameters is often unstable over time. That is, although it has not typically been emphasized as such, tuning curves computed during preparation typically differ from tuning curves during the movement (Crammond and Kalaska, 2000; Wise et al., 1986), and tuning often changes drastically and even multiple times during the movement itself (Churchland and Shenoy, 2007b; Fu et al., 1995; Hatopoulos et al., 2007). Perhaps unsurprisingly, then, careful attempts to distinguish between the possible reference frames for representation of movement in preparatory activity have yielded ambiguous results (Ajemian et al., 2008; Kakei et al., 1999; Scott and Kalaska, 1997; Sergio et al., 2005). Even during movement, the common assumption that neural activity represents something like hand velocity (e.g., Moran and Schwartz, 1999; Wang et al., 2007) fits poorly: neural activity often starts well before and ends well after the movement (e.g., Churchland and Shenoy, 2007b), and the distribution of ‘lags’ in the best fits between neural data and hand velocity are disconcertingly wide (Moran and Schwartz, 1999; Schwartz and Moran, 1999; Wang et al., 2007).

Attempts to interpret motor cortical activity as relating directly to the muscles has faced similar challenges. It has long been known that some neurons are more ‘muscle-like’ than ‘movement-like’ (e.g., Evarts, 1968; Yanai et al., 2008), but this describes only some neurons well (e.g., Churchland et al., 2006b; Kurtzer et al., 2005; Scott, 2008) and leaves totally unexplained how preparatory activity can exist without
causing movement (Green and Kalaska, 2011); some kind of nonlinear ‘gate’ would have to be assumed (Benjamin et al., 2010; Bullock and Grossberg, 1988; Cisek, 2006a).

One problem with this research may be that it relies on making neural recordings during performance of a relatively narrow portion of the monkey’s behavioral repertoire: short, straight reaches. Recordings from a wider swath of motor and premotor cortex during a more comprehensive range of natural movements has revealed that several-millimeter-wide patches of cortex seem to be most responsive during complex, coordinated movements, such as eating or bilateral defensive gestures (Graziano et al., 2002). Moreover, when long-pulse-train microstimulation is performed in these areas, complex movements are evoked (Graziano et al., 2002) that are similar to those nearby neurons prefer (Aflalo and Graziano, 2006a). It has therefore been proposed that the complex local connectivity of motor cortex makes M1 more than simply ‘cables to the spinal cord’ (Graziano, 2011), and that instead these neurons operate as part of a set of behaviorally-relevant primitives (Aflalo and Graziano, 2006b; Graziano and Aflalo, 2007). Concordant with this idea, even when a motor cortical neuron has a reliable, short-latency effect on a muscle, the neuron’s firing rate often correlates poorly with the muscle’s activity pattern during real movements (Schieber and Rivlis, 2007) and even this seemingly close spike-to-spike connectivity is present only in some behavioral contexts (Davidson et al., 2007).

This natural behavior-plus-microstimulation approach has been quite revealing about the relationship of motor and premotor cortex to a wider range of movements. However, it does not appear adequate to explain the space of reaching movements. Specifically, it does not seem to account for the multiphasic nature of neural responses even during brief movements (Churchland and Shenoy, 2007b; Fu et al., 1995; Hatsopoulos et al., 2007), tuning for speed (Churchland et al., 2006b), or the existence of preparatory activity.

Fortunately, there is an alternative to the representational framework: the more mechanistic idea that the job of motor cortex is not to represent movement parameters, but to control them (Churchland and Shenoy, 2007b; Fetz, 1992; Robinson, 1992; Scott, 2008; Todorov and Jordan, 2002; Todorov, 2004). In this view, the correlations
observed between neural activity and movement parameters are incidental instead of fundamental (Mussa-Ivaldi, 1988; Sanger, 1994; Todorov, 2000). By assuming that the controller should be near-optimal according to some set of criteria, the (optimal) control framework has led to numerous behavioral predictions. In particular, by assuming that the controller must contend with noise and wishes to minimize expended energy or wear-and-tear on joints, properties of movement such as the well-known bell-shaped velocity profile, covariance patterns of muscles, and patterns of error on repeated trials follow naturally (Todorov, 2000, 2004; van Beers, 2009). In addition, construction of neural network models that implement controllers have led to (somewhat looser) neural predictions. Specifically, such models predict that most neurons should have complex responses that nonetheless correlate with movement parameters, and should lie in intermediate reference frames, as observed. While these results hardly constitute an ironclad test of control optimality, it is clear that considering the motor system as a controller is compatible with known results and presents a more mechanistic way of viewing the system than simply correlating neural responses with movement parameters.

Taking this viewpoint leads to a new set of essential questions. Most critically, we must know: what are the operating principles of the controller? And, how does the neural circuit implement these principles? That is, if motor cortex is a machine for producing movement, how does the machine work?

To answer such mechanism-oriented questions, we can borrow ideas from dynamical systems analysis to provide techniques and formalisms for describing the rules that govern the motor system. Just as the name implies, this framework aims to elucidate the dynamics of a system. In other words, given the state of the controller at one moment in time, it aims to determine the laws that govern what it will do a few tens or hundreds of milliseconds later.

To define the neural state, we will assume that the relevant variables are the neurons’ firing rates. The state of the system could then be specified as an $n$-dimensional vector containing the firing rates of the $n$ neurons. However, it is widely acknowledged that the firing rates of different neurons are correlated. This uncontroversial observation means that knowing the firing rate of one neuron is often informative.
about the firing rates of others. Once these relationships are known, one would not need \( n \) different numbers to have an excellent guess at all the firing rates; one could instead use a (potentially much) lower number, \( k \). The system is therefore said to be \( k \)-dimensional, and the process of determining the relationship between the \( k \) factors and the \( n \) observed variables is referred to as dimensionality reduction. In M1 and PMd, the overall dimensionality has been estimated as being between approximately 8 and 40, depending on exactly how it is measured (Churchland and Shenoy, 2007b; Yu et al., 2009).

Over the past several years, this dynamical systems perspective has led to a series of experiments from the Shenoy lab, reviewed in detail recently (Shenoy et al., 2011). These experiments followed from the idea that, when preparing for any given movement, there should be a small cluster of possible neural states that constitute the ideal motor plan. This hypothesis was thus deemed the ‘optimal subspace hypothesis.’ While it remained somewhat unclear why the optimal subspace for a given reach was in fact beneficial for making the subsequent movement, this hypothesis nonetheless made a number of predictions. First, it predicted that across-trial state variability may be permitted during baseline, when no movement in particular is being prepared, but that the neural state should be brought as tightly as possible within the optimal subspace during preparation. Indeed, across-trial variability declines from baseline to movement preparation, as assessed by the Fano Factor (Churchland et al., 2006c, 2010a). Second, it predicted that trials on which the neural state was brought especially close to the center of the optimal subspace should be ‘better prepared’ and thus have faster RTs than trials where the neural state was less well controlled and ‘missed’ the optimal subspace. This too held (Churchland et al., 2006c). Third, it correctly predicted that perturbing preparatory activity around the time of the Go cue should erase the RT benefit of having prepared (Churchland and Shenoy, 2007a). Fourth, it predicted that the precise state of the motor plan should impact parameters of the subsequent movement. This relationship was observed in the peak speed of movement (Churchland et al., 2006a). In summary, it is clear that the preparatory state is important for producing the movement. Yet, it has remained opaque exactly how the preparatory state relates to neural activity during movement.
Finally, it is worth noting that most of the above analyses required pooling across many trials, as is nearly universal in systems neuroscience. However, there are numerous cases where averaging over many trials relates at best an incomplete story. In general, there are many possible sets of dynamics that are difficult to distinguish without observing single-trial neural trajectories, such as attractor dynamics vs. integration to a bound (Churchland et al., 2007, 2011; Yu et al., 2009) (Figure 1.3B,D), or whether some dimensions have different dynamics than others (Figure 1.3C).

Moreover, some processes are simply intrinsically single-trial, such as decision-making (Figure 1.3A). When making decisions, we expect the neural state over time – the neural ‘trajectory’ – to have a different temporal evolution and potentially even a completely different course from trial to trial. With a single-trial view, it should be possible to determine, for instance, what ‘went wrong’ on a slow RT trial (Churchland et al., 2010b) or error trial (Cisek and Kalaska, 2005); why some trials have faster RTs than others (Afshar et al., in press); and what happens during ‘changes of mind’ and whether they occur spontaneously (Resulaj et al., 2009).

1.3 Summary of thesis work

In this thesis, I will address the relationship of preparatory activity to movement activity, the nature of movement activity, and single-trial dynamics during motor preparation in the context of decision-making. My overarching finding is that it is difficult to understand the responses of individual neurons; instead, one must consider the population of neurons together to make sense of the system’s operating principles.

Chapters 2 and 3 will focus on the basic question of why movement activity causes movement and preparatory activity does not. Two datasets from different monkeys are presented for PMd and two more for M1. Using the waveform shape of the action potential, putative inhibitory neurons are distinguished from excitatory neurons. It will be argued that, contrary to common assumption, inhibitory neurons do not appear to be ‘gating’ excitatory outputs to the spinal cord during motor preparation in either PMd or M1. Instead, it appears that a rather different mechanism is involved in preventing premature movement. Specifically, population analyses will demonstrate
Figure 1.3: Conceptual illustration showing how single-trial analytical methods can be applied to different behavioral tasks, including those involving perception, decision making, attention, and motor planning. The neural mechanisms underlying these behavioral tasks may involve (A) switching between two possible percepts or decisions, (B) rising to threshold, (C) decaying along a single slow mode, (D) or converging to an attractor. Each panel includes icons of the relevant behavioral tasks and brain areas (top), single-trial neural trajectories in the firing-rate space of 2 neurons (bottom left), and corresponding firing-rate profiles (both single-trial and trial-averaged) for each neuron (bottom right). Task icons represent random dot motion discrimination, binocular rivalry, ambiguous motion, vibrotactile match-to-sample, and delayed-match reaching (A); saccade initiation and motion detection (B); go-nogo in the presence of distractors (C); and simple delayed reaching (D). From Yu et al. (2009).
that neurons compensate for one another during preparation, remaining in an ‘iso-
force space.’ In other words, the population patterns its outputs to maintain a net
zero impact on the muscles.

Having considered the key difference between preparatory and movement activity,
Chapter 4 will explore how they are related. This chapter will argue that prepara-
tory activity actually acts to seed the dynamics that prevail during movement. At
the single-neuron level, these dynamics yield complex, confusing responses. However,
taken as a population, a novel analysis method and data from five monkeys will show
that the dynamics are in fact dominated by straightforward oscillations.

In pursuit of causal perturbations to complement the observational analyses above,
Chapter 5 presents technical advances in bringing optogenetic techniques to monkey.
Successful application of excitatory, inhibitory, and special-purpose opsins will be
demonstrated. In addition, it will be shown that optogenetic stimulation does not
appear to have the same effect on motor cortex as electrical microstimulation or
optogenetic stimulation in transgenic mice.

The final chapter presents preliminary results from an ongoing project on neural
dynamics during single trials of decision-making. Complete datasets from two mon-
keys are presented. Leveraging new techniques for extracting low-dimensional neural
trajectories from simultaneous neural recordings, we will show that the motor cortex
reflects motor decision-making at a fine timescale. In particular, these results will
demonstrate that, in a free-choice paradigm, vacillation between decision options is
present during motor preparation in the monkey.
Chapter 2

Lack of inhibitory gating in PMd

This study was published as “Roles of Monkey Premotor Neuron Classes in Movement Preparation and Execution” in the Journal of Neurophysiology, vol. 104, pp. 799-810, 2010. MTK collected half the data, designed and performed all analyses, and wrote the manuscript.

Dorsal premotor cortex (PMd) is known to be involved in the planning and execution of reaching movements. However, it is not understood how PMd plan activity – often present in the very same neurons that respond during movement – is prevented from itself producing movement. We investigated whether inhibitory interneurons might ‘gate’ output from PMd, by maintaining high levels of inhibition during planning and reducing inhibition during execution. Recently-developed methods permit distinguishing interneurons from pyramidal neurons using extracellular recordings. We extend these methods here for use with chronically-implanted multi-electrode arrays. We then applied these methods to single- and multi-electrode recordings in PMd of two monkeys performing delayed-reach tasks. Responses of putative interneurons were not generally in agreement with the hypothesis that they act to gate output from the area: in particular it was not the case that interneurons tended to reduce their firing rates around the time of movement. In fact, interneurons increased their rates more than putative pyramidal neurons during both the planning and movement epochs. The two classes of neurons also differed in a number of other ways, including
greater modulation across conditions for interneurons, and interneurons more frequently exhibiting increases in firing rate during movement planning and execution. These findings provide novel information about the greater responsiveness of putative PMd interneurons in motor planning and execution, and suggest that we may need to consider new possibilities for how planning activity is structured such that it does not itself produce movement.

2.1 Introduction

Dorsal premotor cortex (PMd) is known to be involved in both the planning and execution of reaching movements. Frequently, both planning and movement-related activity are exhibited in the very same neurons (Tanji and Evarts, 1976; Weinrich and Wise, 1982). We therefore ask a basic question: why does PMd plan activity not drive movement?

When a monkey is cued about the path of an upcoming reach but required to withhold it until a go cue, PMd activity exhibits tuning for parameters of the reach during the plan period (Churchland et al., 2006b; Godschalk et al., 1985; Hocherman and Wise, 1991; Messier and Kalaska, 2000; Riehle and Requin, 1989). PMd activity also predicts reaction time (RT; Churchland et al., 2006c; Riehle and Requin, 1993) and variability in the upcoming movement (Churchland et al., 2006a). Further arguing that PMd is specifically involved in movement planning, disruption of PMd activity near the time of the go cue delays movement onset (Churchland and Shenoy, 2007a). Additionally, it is known that PMd sends projections down the spinal cord to motor interneurons (Dum and Strick, 1991), and microstimulation in PMd causes arm movements (e.g., Weinrich and Wise, 1982). Hence, movement-epoch PMd activity presumably contributes to actually driving the movement. Theoretical treatments thus often assume a ‘gate’ between plan and movement-related neurons (e.g., Bullock and Grossberg, 1988; Cisek, 2006a), and pharmacological experiments have suggested that reduction of inhibition might impair the ability to withhold premature movements (Sawaguchi et al., 1996).

In the oculomotor system, this gating process is fairly well understood. During
saccade preparation, cortical areas such as the frontal eye fields (FEF) generate a representation of the upcoming saccade. A downstream subcortical area (nucleus raphe interpositus) contains ‘omnipause neurons’ (Cohen and Henn, 1972), which tonically inhibit saccade-generating ‘burst neurons’ in the brainstem (paramedian pontine reticular formation and rostral interstitial nucleus of the median longitudinal fasciculus; for review, Fuchs et al., 1985). Omnipause neurons stop firing during the saccade, releasing their inhibitory gate of burst neuron output (Evinger et al., 1982). Since there is no brainstem intermediate between cortex and the arm muscles, and inhibition in cortex is almost exclusively local (Markram et al., 2004), a hypothetical mapping of this mechanism to PMd is shown schematically in Figure 2.1. In this mapping, corticospinal neurons might correspond to burst neurons, local interneurons to omnipause neurons, and local pyramidal neurons to upstream areas such as FEF.

While PMd sends axons down the spinal cord, analogously to burst neurons, it is a cortical area, like FEF. There is also a small amount of preparatory activity even in the spinal cord (Prut and Fetz, 1999). Thus, it is not clear whether PMd should be more closely analogous to FEF or brainstem saccade areas, or use some other mechanism entirely to prevent premature movements. It is known that PMd can exert inhibitory as well as excitatory effects on primary motor cortex (M1; Ghosh and Porter, 1988; Keller and Asanuma, 1993; Tokuno and Nambu, 2000), which might point towards a feedforward-inhibition gating mechanism.

A recent study examined whether an oculomotor-like output-gating mechanism might be at play in forelimb movements in rats (Isomura et al., 2009). They found evidence against such a mechanism, but they also note that rats do not have a clear PMd-M1 separation and found that interneurons were only weakly tuned, in contrast to known interneuron tuning in monkey M1 (Merchant et al., 2008). Thus, it remains unclear what mechanism is at work in preventing PMd plan activity from driving movement in the monkey.

Different hypotheses about movement gating make different predictions regarding the relative activity patterns of interneurons and pyramidal neurons. Examining differences in their patterns of activity could thus be informative regarding the gating mechanism. We therefore wished to identify interneurons and pyramidal neurons
in our recordings. Since interneurons have briefer action potentials than pyramidal neurons (Connors and Gutnick, 1990; McCormick et al., 1985), and the extracellular waveform reflects the intracellular waveform (Henze et al., 2000), the extracellularly-recorded waveform duration can be used to distinguish interneurons from pyramidal neurons with substantial reliability (Bartho et al., 2004). This technique has previously been used in primary somatosensory cortex (Simons, 1978; Swadlow, 2003), prefrontal cortex (Diester and Nieder, 2008; Johnston et al., 2009; Rao et al., 1999; Wilson et al., 1994), V4 (Mitchell et al., 2007), FEF (Cohen et al., 2009), and M1 (Merchant et al., 2008).
We first replicate prior findings of a bimodal distribution of spike-waveform durations for single-electrode recordings. We then extend this method for use with chronically-implanted electrode arrays. We observed consistent differences in the activity patterns of putative interneurons versus putative pyramidal neurons. Those differences did not follow the predictions of the output-gating ‘omnipause’ hypothesis; interneuron activity was highest, not lowest, around movement onset. Nor were our results consistent with other, more refined variants of the output-gating hypothesis. These results imply that either gating is not accomplished through inhibition, or that the gate is located downstream of PMd.

2.2 Methods

2.2.1 Subjects

Animal protocols were approved by the Stanford University Institutional Animal Care and Use Committee. Subjects were two adult male macaque monkeys (Macaca mulatta) trained to perform variants of the delayed reach task for juice reward. After initial training, we performed a sterile surgery during which the monkeys were implanted with a head restraint and either a 96 electrode silicon array (monkey H) or a standard recording cylinder (monkey J). The electrode array (Blackrock Microsystems, Salt Lake City, UT) was implanted in caudal PMd (adjacent to primary motor cortex), as estimated visually from local anatomical landmarks (Figure 2.2A). Array recordings yielded strong, well-tuned plan-period activity for arm movements.

The cylinder (Crist Instruments, Hagerstown, MD) was centered over caudal PMd, initially estimated using stereotaxic coordinates (13-17 mm anterior to stereotaxic zero, the intermeatal ear bar line) and from previous surgeries and MRIs in other monkeys. The cylinder was placed surface normal to the skull, which was left intact and covered with a thin layer of dental acrylic. To accommodate recording, 3 mm holes were drilled later under ketamine/xylazine anesthesia.

We confirmed the location of our cylinder with a subsequent craniotomy (Figure 2.2B) as part of a later array implantation surgery. The single-electrode recordings
Figure 2.2: Relative locations of narrow- and broad-spiking neuron recordings. (A) Location of the implanted electrode array in monkey H. Nearby sulci are shown with black lines. Red dots indicate narrow-spiking cells, blue dots indicate broad-spiking neurons. Note that some electrodes recorded two isolable neurons, indicated by two dots in the box for that electrode. (B) Approximate location of the region of single-electrode recordings in monkey J. (C) Expanded view of single-electrode penetration locations for monkey J. Blue and red are as in panel A, black dots indicate non-canonical waveform neurons. Dots are scattered <0.1 mm to reveal overlapping recordings. (D) Counts of each neuron type by depth for monkey J. Red and blue are as in other panels; gray indicates neurons with non-canonical waveforms. Neurons unclassifiable due to flat peaks not shown.
reported here appear to be in PMd proper just lateral of the precentral dimple, though potentially near the PMd/M1 ‘transition zone’ (Keller, 1993; Weinrich and Wise, 1982; Wise et al., 1986). While we cannot definitively localize PMd without histology, consistent with the PMd classification plan-period activity was common and robust for PMd recordings, but not for surface or sulcal M1 sites (recorded separately, not otherwise reported here). Also, microstimulation thresholds for our PMd recordings were higher (median 50 µA) than in M1 (median 25 µA). Microstimulation evoked movements of the shoulder and upper arm, or (much less often) of the wrist. A number of our recordings were made deeper (>4 mm) than the typical depth of cortex, consistent with the thickening of cortex near the precentral dimple.

2.2.2 Task apparatus

We used the same task apparatus as described previously (Churchland et al., 2006c). Briefly, during experiments monkeys sat in a customized chair (Crist Instruments) with the head restrained. The left arm of monkey J was restrained loosely using a tube and a cloth sling; monkey H did not have an arm restraint, but seldom moved the non-reaching arm from his side. Stimuli were back projected onto a frontoparallel screen ∼27 cm from the eyes (the exact distance depended on the size of the monkey). A photodiode was used to record the timing of video frames with 1 ms resolution. The position of a reflector taped to the fingers was tracked optically in the infrared (Polaris system; Northern Digital, Waterloo, Ontario, Canada). The eyes were also tracked in the infrared (Iscan, Burlington, MA). A clear acrylic shield prevented the monkey from touching the dichroic eye-tracking mirror or from bringing the reflector to his mouth. A tube fixed to this shield dispensed juice rewards.

2.2.3 Task design

The tasks for both monkeys were variants of the center-out delayed reach task (Figure 2.3), described previously (Churchland et al., 2006c). Experiments consisted of trials, each a few seconds long, which ended in a juice reward if successful. The animal began a trial by fixating and touching (for at least 400 ms) a fixation spot, after which a
Figure 2.3: Delayed reach tasks. Each row illustrates a single trial, taking 2-3 seconds. Top row illustrates task for monkey H, bottom row for monkey J. Monkey H touched targets directly with his fingers; monkey J also touched the screen but contacted the targets with a cursor floating 2.5 cm above his hand. The animals first fixated (eye and hand/cursor) fixation cues (>400 ms), after which a target appeared. After a plan period (H: 200-1000 ms; J: 0-1000 ms), the go cue was given. For monkey H, the go cue was indicated by a slight enlarging of the target; for monkey J, the target jittered slightly during the plan period and the cessation of jittering indicated “go.” In both cases the fixation point was also extinguished at the time of the go cue. For monkey J, most trials also involved barriers which appeared at the same time as the target, instructing a curved reach.

target appeared. After a plan period, a go cue was given, and reaches were rewarded if they were brisk and accurate. Reward was delivered after the target was held for 300 ms (monkey H) or 450 ms (monkey J), with the next trial beginning a few hundred milliseconds later.

For monkey H, the fixation spot was central and the plan period was 200-1000 ms. Only trials with delay periods >400 ms were analyzed. The go cue was a slight enlargement of the target and the disappearance of the fixation spot. RTs were required to fall between 150 and 500 ms. Two rings of 8 targets were used, for a total of 16 conditions. For this monkey, each target location is hereafter referred to as a ‘condition.’

Monkey J performed a novel variant of the delayed-reach paradigm, called the maze task. This task also required delayed reaches (0-900 ms plan period) to targets. Again, only trials with delay periods >400 ms were analyzed. The maze task includes
a number of additional complexities (explained below), but here we use it simply as a 24-condition delayed reach task. In the maze task the monkey touched the screen, but contacted the targets with a virtual cursor floating 2.5 cm above his hand. Additionally, we required that the cursor path not pass through a set of virtual barriers which varied trial to trial. Thus, the monkey was required to make curved reaches on most trials. Cursor contact with a barrier resulted in an aborted (un-rewarded) trial. Barriers appeared at the same time as the target, and trials with barriers were interleaved with no-barrier trials. Mazes varied in start point, end point, and barrier positions. Different mazes were used for different neurons, with 24 stimuli (including those with and without barriers) for each neuron. Effort was made to span the major reach directions and curve shapes for each neuron recorded. The target jittered slightly (2 mm) when it appeared, and the cessation of this jitter and disappearance of the central fixation spot comprised the go cue. RTs were required to fall between 150 and 600 ms. For this monkey, each unique maze (i.e., start point, end point, and reach shape) is hereafter referred to as a ‘condition.’

### 2.2.4 Neural recordings, classification of neuron types, and EMG

Neural and EMG recordings were made using previously described techniques (Churchland et al., 2006c). Neurons were classified as narrow- or broad-spiking using waveform analysis techniques similar to those in Mitchell et al. (2007). Full details are given in the Supplementary Methods.

### 2.2.5 Neural analyses

We operationally define each neuron’s ‘cross-condition modulation pattern’ as simply its mean response (in some time window) across conditions (16 or 24 conditions, depending on the monkey). This is effectively a tuning pattern, but no parametric tuning model (e.g., for direction, endpoint, etc.) is assumed. To obtain the cross-condition modulation pattern, we took the mean firing rate over the relevant epoch for each condition, resulting in a 16- or 24-element vector. We define the ‘cross-condition
modulation’ as the maximum element of the vector minus the minimum element. For analyses of tuning consistency, we produce one such vector for the baseline period, one for the plan period, and another for the peri-movement period. We subtract the baseline vector from each the plan and peri-movement vectors, and correlate the two resulting vectors.

2.3 Results

2.3.1 Separation of narrow- and broad-spiking neurons

Previous work has established the trough-to-peak duration (TTP) of the spike waveform as among the more reliable methods for distinguishing putative interneurons from pyramidal neurons (Bartho et al., 2004; Mitchell et al., 2007). Specifically, the distribution of TTPs is often bimodal, with the briefer mode thought to correspond with inhibitory interneurons and the second mode corresponding predominately with pyramidal neurons (Krimer et al., 2005). The only known mammalian narrow-spiking neurons that are excitatory, spiny stellate cells, are not thought to be common outside of Layer IV of primary sensory cortex (Okhotin, 2006). Additionally, previous studies in M1 and elsewhere have used spike-triggered averages to demonstrate that neurons with narrow spike waveforms are inhibitory with substantial reliability (Bartho et al., 2004; Merchant et al., 2008). Nonetheless, since we cannot know for certain that every neuron with a narrow waveform is an interneuron, or that every neuron with a broad waveform is not, we refer to them as narrow-spiking neurons and broad-spiking neurons to be explicit that we are not directly measuring whether each neuron is inhibitory or excitatory.

In monkey J, we recorded 163 neurons using single electrodes. We replicated the bimodal distribution of TTPs (Figure 2.4A) found by Mitchell et al. (2007). The bimodality of our TTP distribution was significant with $p = 0.011$ using Hartigan’s dip test with a bootstrap (Hartigan and Hartigan, 1985; Mechler and Ringach, 2002).

For monkey J, 34 neurons were identified as narrow-spiking (putative interneurons) and 79 were identified as broad-spiking (putative pyramidal neurons). We excluded
CHAPTER 2. LACK OF INHIBITORY GATING IN PMD

from analysis 14 neurons with flat-topped after-potentials whose peak time could not be reliably measured. An additional 36 had non-canonical waveforms (e.g., missing a post-trough peak); these neurons are analyzed separately in a subsequent section. At most of the recorded depths, a consistent proportion of narrow-spiking neurons was found, though deeper recordings yielded mostly non-canonical waveforms (Figure 2.4D). The fraction of narrow-spiking neurons found, 30% of identified neurons (21% of all recorded neurons), is consistent with the proportion found by others (26%: Merchant et al. 2008; 27%: Mitchell et al. 2007; 27%: Rao et al. 1999) and with the estimated fraction of interneurons in cortex, 20-30% (Connors and Gutnick, 1990).

In monkey H, we recorded 71 single units using a chronically-implanted multi-electrode array. 20 neurons were identified as narrow-spiking and 31 as broad-spiking (Figure 2.4B). This corresponds to 39% of identified neurons being narrow-spiking.

Figure 2.4: (A) Histogram of the trough-to-peak (TTP) duration for monkey J. Inset shows amplitude-normalized mean waveforms. Red indicates narrow-spiking neurons, blue indicates broad-spiking neurons. (B) Same for monkey H using simulated MAP filters on broadband data from the chronic electrode array. Red and blue are as above, green indicates neurons considered unclassifiable. Special steps were taken to address the lack of two clear peaks in this distribution (see Supplementary Methods), and subsequent analyses yield very similar results for both the single-electrode data and the electrode array data.

A

B
This modest over-representation of narrow-spiking neurons may occur because multiple broad-spiking neurons were commonly present on a single channel of the array recordings. High-quality spike sorts were thus achievable for a greater fraction of the (relatively isolated) narrow-spiking neurons, and they were therefore included at a higher rate.

The precise depths of the implanted array tips are not known, but the electrodes are 1 mm long and were fully inserted immediately after surgical implantation. The electrode tips were therefore likely shallower on average than our single-electrode recordings. Thus, the array recordings were presumably primarily from the shallower cortical layers, while the single-electrode recordings sampled all layers. To our knowledge, narrow- and broad-spiking neurons have not previously been identified using this recording technology. Although special steps were required to address the lack of two clear peaks in the TTP distribution (see Supplementary Methods), subsequent analyses yield very similar results for both our single-electrode data and our electrode array data.

### 2.3.2 Comparison of the responsiveness of the two neuron classes

Figure 2.5 shows example PSTHs for one broad-spiking neuron, one narrow-spiking neuron, and one deep non-canonical neuron (to be discussed later). Though there was wide variation in the response patterns of different neurons within each class, these examples exhibit the most common patterns recorded. During the plan period, the broad-spiking neuron (blue) shows increases in firing rate for some conditions and decreases for others. The narrow-spiking neuron (red) is also strongly tuned during the plan period, but most of its firing rate changes are positive. For this neuron, firing rates rise even further shortly before movement onset, and drop back to baseline several hundred milliseconds after movement onset.

Figure 2.6 (top row) plots population PSTHs for narrow-spiking neurons and for broad-spiking neurons, collapsing over all conditions. There are substantial average differences between narrow- and broad-spiking neurons. Narrow-spiking neurons tend
to increase their firing rates more strongly during the plan period, and to show a much greater rise in firing rate before movement onset (as in the examples). Firing rates do not return to baseline until after movement onset. Surprisingly, the mean rate for broad-spiking neurons is almost unchanged between the baseline period and the plan period. That is, on average pyramidal neurons do not fire at higher rates during motor planning than at baseline. The flatness of the aggregate PSTH for broad-spiking neurons does not imply that they have weak cross-condition modulation, however. Instead, it indicates that on average their firing rates decrease as often as increase, as in the example neuron (Figure 2.5, left).

The bottom row of Figure 2.6 plots mean cross-condition modulation, defined as the range of firing rates over all conditions at each time point independently. For both narrow- and broad-spiking neurons, cross-condition modulation rises substantially at target onset and stays high through movement onset. However, cross-condition modulation is greater for narrow-spiking neurons both during the plan period and around
Figure 2.6: Comparison of response properties between narrow- and broad-spiking neurons. Changes in firing rate for monkey J (left); monkey H (right). *Top row*: mean firing rate across conditions. Red trace shows the average firing rate across all narrow-spiking neurons across all conditions (both preferred and non-preferred). Blue trace shows the same for broad-spiking neurons. These plots are essentially population PSTHs. *Bottom row*: Red trace shows the cross-condition modulation (most preferred minus least preferred condition at every time point) for narrow-spiking neurons. Blue trace is the same for broad-spiking neurons. These plots are effectively a non-parametric ‘depth of tuning’ over time. Note that treating time points independently causes baseline cross-condition modulation to be greater than zero due to noise. TARGET is target onset time; MOVE is movement onset time. Flanking traces are SEMs across neurons.
movement onset. Together with the mean PSTHs, this means that both classes of neurons show substantial cross-condition modulation, but cross-condition modulation was stronger for narrow-spiking neurons and their firing rates were more likely to increase than decrease. This contrasts sharply with results from rat forelimb motor cortex, in which interneurons show cross-condition modulation almost exclusively during the movement epoch, and show little tuning for movement direction (Isomura et al., 2009). Merchant et al. (2008), however, found a generally similar pattern of firing rate changes in monkey primary motor cortex.

To see whether the structure of the population PSTHs was representative, we performed a cell-by-cell analysis. For each neuron, we took the mean plan-period firing rate (50 to 400 ms after target onset) minus the mean fixation-period baseline firing rate. The histogram over neurons is plotted in the top row of Figure 2.7. Consistent with the population PSTHs, the distribution for narrow-spiking neurons is shifted slightly to the right of that for broad-spiking neurons in both monkeys (monkey J, narrow vs. broad medians: 4.9 vs. 0.8 spikes/s, p < 0.02; monkey H: 0.5 vs. 0 spikes/s, p < 0.001, Mann-Whitney U tests). The second row plots movement-period activity (-100 to +200 ms from movement onset) minus the baseline, and here the rightward shift for narrow-spiking neurons is more pronounced (J: 12.6 vs. 0.5 spikes/s, p < 0.002; H: 5.8 vs. 1.3 spikes/s, p < 0.03), again consistent with the population PSTHs. Results were similar when the analysis was performed over neuron-conditions instead of neurons.

We also tested whether neurons exhibited a ‘pause-like’ pattern of activity, as was hypothesized for interneurons in the output-gating model (see 2.1B). We took the mean plan-period activity minus the mean movement-period activity for each neuron, and plot the distributions in the bottom row of Figure 2.7. A neuron with a tendency to be very active during plan and only weakly active during the movement has a large value for this measure. No significant difference was observed between cells classes in either monkey for this metric, and the trends present were not in the hypothesized direction (J: -3.6 vs. 0.8 spikes/s, p > 0.14; H: -5.3 vs. -0.8 spikes/s, p > 0.07). Thus, it was not the case that one class of neuron tended to have a high firing rate during the plan period and pause during the peri-movement period. Additionally, note that
Figure 2.7: Cell-by-cell analysis of firing rates. Distributions for monkey J (left); monkey H (right). This analysis summarizes patterns from 2.6 for individual neurons. 

Top row: For each neuron, the mean plan-period firing rate (FR) is taken and the mean baseline-period FR is subtracted. The histogram over neurons is plotted. Red indicates narrow-spiking neurons, blue indicates broad-spiking neurons. Note that the distributions are unimodal and nearly symmetrical, and that the distribution for narrow-spiking neurons is shifted to the right of that for broad-spiking neurons. That is, narrow-spiking neurons tended to have higher FRs during planning than during baseline. This shift was significant for both monkeys (see Results).

Middle row: Same for mean movement-period FR minus mean baseline-period FR. Note similar shift as for the plan period above. This shift was significant for both monkeys. Bottom row: mean plan-period FR minus mean movement-period FR. High values indicate a pause-like neuron. The broad- and narrow-spiking distributions were not significantly different for either monkey.
these distributions appear to be unimodal; there does not appear to be a subset of ‘pause-like’ neurons. Results were again similar when the analysis was performed over neuron-conditions instead of neurons.

From the output-gating hypothesis illustrated in Figure 2.1, one might have expected a particular pattern of activity from interneurons: high and perhaps un-tuned firing rates during the plan period, and a pause in firing around movement onset. This is not what was observed for the narrow-spiking neurons (putative interneurons). Instead, such neurons generally showed strong cross-condition modulation during the plan period, followed by a rise in activity around movement onset. Nor was there a clear subset of narrow-spiking cells with a pause-like response. These patterns do not appear consistent with a straightforward formulation of the output-gating hypothesis, for which inhibition should decline around movement onset.

2.3.3 Additional controls for neuron classification

We considered whether the observed effects could be a result of neuron-selection bias. For the single electrode recordings, we attempted to isolate every neuron encountered. Recordings were abandoned only if the isolation was lost, the neuron died, or the neuron was completely unresponsive during the plan period. Few neurons were abandoned due to a lack of responsiveness, perhaps because the novel task used with monkey J evoked particularly strong responses. While our selection criteria could lead to a general bias towards responsiveness, they should not lead to a systematic difference between the classes. Further, since we recorded as many narrow-spiking neurons as expected from the percentage of interneurons in cortex, our criteria probably did not strongly bias us towards or away from recording interneurons. For the array recordings, neuron selection was made purely on quality of isolation. Again, this should not produce response differences between neuron classes, and similar results were found with the array recordings as with single electrodes.

We also tested whether our effects were sensitive to the precise threshold for separating narrow- and broad-spiking neurons. We re-divided the population using a range of thresholds from 170-300 µs and calculated the mean difference between
the population PSTHs for the two classes. The differences varied smoothly with the threshold value in both monkeys. Thus, the observed difference in firing rate changes is largely insensitive to the precise threshold value.

2.3.4 Consistency of neural modulation across conditions between planning and movement

We also considered that inhibitory interneurons might change their preferred condition between the plan period and the movement period (as observed by Rao et al. (1999) in prefrontal cortex). During planning of a movement, one set of interneurons could prevent premature execution while pyramidal neurons achieved the needed pattern of rates. During execution of the movement, these interneurons might decrease their inhibition, while another set of interneurons inhibited competing movements. Instead of an overall pause in firing for interneurons, this version of the output-gating hypothesis predicts that interneurons would invert their preferred conditions between the plan and movement epochs. In contrast, pyramidal neurons would be expected to maintain consistent preferred conditions. Thus, what was inhibited during plan would be released during movement.

We did not find this pattern in our data. We frequently observed changes in the preferred condition between the plan and movement periods. However, narrow-spiking neurons did not tend to fully invert their preferred conditions, and such preference changes were just as prevalent for broad-spiking neurons as for narrow-spiking neurons. PSTHs for three example neurons, all broad-spiking, are shown in Figure 2.8 (top). These neurons are all well-modulated across conditions at nearly every time point. One of these neurons had consistent preferred conditions between the plan and movement periods (Figure 2.8, top right). The other two examples’ patterns of cross-condition modulation change dramatically between the plan period and the peri-movement period: some most-preferred conditions during planning became least-preferred during movement, and vice versa. Such shifts are not visible in the population plots (Figure 2.6), where cross-condition modulation was computed as the most-preferred minus least-preferred condition at each time point independently.
Figure 2.8: Relationship between plan and peri-movement cross-condition modulation patterns. Top: PSTHs for three example neurons, all broad-spiking (monkey J P59, P57, P23). For these examples, the correlations between the plan and peri-movement cross-condition modulation patterns were -0.53 (left), 0.17 (center), and 0.86 (right). Colors correspond to different conditions. Conditions selected are representative; not all are shown. Bottom: Distribution of correlation values. A near-zero correlation means that the neuron showed little relationship between its plan and peri-movement modulation patterns. A negative value implies inverted preferences during the two epochs. Left bottom: Histogram of correlations for narrow-spiking neurons; right bottom: for broad-spiking neurons. Neurons are pooled across monkeys: monkey J (H) contributed 33 (8) narrow-spiking neurons and 79 (4) broad-spiking neurons (note that many of monkey H’s neurons had shallow cross-condition modulation for one epoch and were thus excluded from this particular analysis; see Results). Black arrows are distribution means. For narrow-spiking neurons, the mean for monkey J (H) is 0.26 (0.16), and for broad-spiking neurons the mean is 0.17 (0.09). Gray arrows are bootstraps (see Results).
To quantify the extent of such changes in cross-condition modulation, and to ask whether changes were more common for narrow-spiking neurons, we determined the consistency between plan modulation and movement modulation for each cell. For each well-modulated neuron (>5 spikes/s cross-condition modulation for plan and move), we computed its cross-condition modulation pattern: a vector containing the mean rate for each condition over an epoch, minus the baseline firing rate for that condition. We then correlated the cross-condition modulation patterns for the plan period (50 to 400 ms after target onset) with those from the peri-movement period (-100 to +200 ms from movement onset). These correlations spanned a broad range (Figure 2.8, bottom): being sometimes near one (very similar modulation patterns during planning and movement) and sometimes near negative one (pattern of preferences inverted). On average, correlations were only slightly above zero (little relationship between the two patterns, neither consistent nor inverting). This was true for both monkeys for both narrow- and broad-spiking neurons. These weak correlations using cross-condition modulation patterns is consistent with previous findings using a cosine-tuning model (Crammond and Kalaska, 2000), with population statistics (Johnson et al., 1999), and with oscillatory premotor neurons (Lebedev and Wise, 2000). Results when fitting our data with cosine tuning curves yielded similar patterns between narrow- and broad-spiking neurons as well (Figure 2.9).

As a control, we considered whether this overall low correlation between plan- and movement-period modulation patterns could be due to noise: if we poorly estimated the mean firing rate due to low trial counts, this could reduce the correlation between patterns that were actually similar. To estimate the expected departure from unity correlation due to sampling noise, we re-sampled trials and calculated the re-sampled plan (move) modulation pattern against the original plan (move) modulation pattern; these re-samplings yielded values rather close to 1 (Figure 2.8 gray arrows), implying that the wide range of correlations between the plan and move periods are not due to sampling noise. Finally, we considered whether having subtracted baseline activity to compensate for posture tuning might have reduced the correlations. However, when baseline activity was not subtracted, correlations were even lower.
Figure 2.9: Analysis of preferred directions for monkey H (monkey J did not perform a purely center-out task). A preferred direction (PD) was obtained for each neuron by regressing a plane over the points defined by the X- and Y-target positions and the firing rates for the corresponding targets, then taking the plane’s direction of maximal derivative. Neurons were included if their mean firing rate across targets was at least 1 spike/s and a bootstrap of the PD yielded a 95% confidence interval of less than pi/2 radians. Top row: Histograms of PDs during the plan epoch (50 to 400 ms after target onset). Bottom row: Histograms of PDs during the peri-movement epoch (-100 to +200 ms from movement onset). Left: Narrow-spiking neurons. Right: Broad-spiking neurons. There are no obvious differences between the distributions of PDs for narrow- and broad-spiking neurons, nor is there a clear tendency for PDs to rotate in one direction or the other between the plan and peri-movement epochs for either class.

2.3.5 Non-canonical waveforms

In classifying neurons into narrow- and broad-spiking, we excluded waveforms that did not conform to the classical pattern of a trough followed by a peak. In particular, this included many neurons that had only a single positive peak, or a positive peak followed by a small trough and with little post-trough peak (Figure 2.10A). Computer simulations, in combination with systematically positioned extracellular recordings, suggest that such waveforms may be obtained from the distal portion of pyramidal neurons’ dendrites because of capacitative currents (Gold et al., 2006). In order to
isolate a distal dendritic recording with other neurons nearby, presumably the neuron must be large. Since non-canonical waveforms were found almost exclusively deep in cortex (Figure 2.2D), this may indicate that they belong primarily to large pyramidal neurons in layers V and VI, which are apt to project to other brain areas or down the spinal cord. Merchant et al. (2008) have recently argued for an identifiable subgroup of putative pyramidal neurons in primary motor cortex, characterized by having long duration waveforms, having seemingly different connectivity, and being located primarily in Layer V. We suspect that the neurons we identify as having non-canonical waveforms may be a similar subgroup as that identified by Merchant et al. (2008), based on their depth. We found non-canonical waveform neurons almost exclusively in recordings from monkey J; the lack of non-canonical waveforms from the multi-electrode arrays is likely due at least in part to their having electrodes 1 mm long, and thus probably not recording from the deeper layers.

We analyzed these neurons separately. The example neuron shown in Figure 2.5 (right) is representative of this neuron class. Such neurons typically had modest changes in firing rate during the plan period, with much greater (and mostly positive) changes in firing rate during the movement. This movement-period activity seems to be maintained relatively later into the movement than for the other two neuron classes. The pattern illustrated in the example neuron can be seen in the

Figure 2.10: Analysis of ‘non-canonical waveform’ neurons (all from monkey J). (A) Normalized waveforms. (B) Mean firing rate across conditions and neurons over time (same format as Figure 2.6 top left). (C) Cross-condition modulation (preferred - least preferred condition, same format as Figure 2.6 bottom left). Black lines show responses of non-canonical neurons. Flanking traces are SEMs.
population PSTH and cross-condition modulation plots for non-canonical waveform neurons (Figure 2.10B, C). This might indicate the presence of a mechanism for limiting the activity of output neurons within PMd.

2.4 Discussion

These results show that one can distinguish, based on extracellularly-recorded waveforms, two classes of neurons in PMd: narrow-spiking (putative interneurons) and broad-spiking (putative pyramidal neurons). These classes differ, as populations, in their patterns of activity during both planning and execution of movements. Compared with putative pyramidal neurons, putative interneurons have stronger cross-condition modulation, are more likely to have firing rates that rise during motor planning, and are more likely to have firing rates that rise around the time of movement onset.

2.4.1 Interneurons are more responsive than pyramidal neurons

The finding that interneurons are more strongly modulated across conditions than pyramidal neurons is consistent with their larger dynamic range (Connors and Gutnick, 1990). More surprising was the finding that putative interneurons are more likely to have firing rates that rise during movement planning, while putative pyramidal neurons showed more symmetric firing rate changes. To put this another way, for interneurons most conditions (most reach directions / reach paths) caused an increase in firing rate, while for pyramidal neurons, conditions that caused suppression were almost as common as those that caused excitation. This is perhaps remarkable - one usually supposes that overall activity goes up during motor planning - yet our results indicated the average rate of a population of pyramidal neurons rises only slightly. Most of the increase in overall rate (and what is perhaps detected with lower-resolution methods such as fMRI) is contributed by interneurons. The functional significance of the rise in inhibition, but not in excitation, is unclear. One speculation
is that as incoming excitation rises, the relative rise in internal inhibition may serve to maintain a balance of excitation and inhibition (Shadlen and Newsome, 1998; van Vreeswijk and Sompolinsky, 1996). Physiologically, the asymmetry may arise because many interneurons have chloride reversal potentials near threshold (Martina et al., 2001). If such neurons are not readily hyperpolarized by GABA, they may be less likely to show declines in firing rate.

2.4.2 Lack of support for output gating by inhibition

Contrary to what is expected if inhibition served an output-gating function for PMd, putative interneurons were actually most active around movement onset. Indeed, of the two cell classes, interneurons showed the larger activity increase at that time. Thus, it does not appear that inhibition falls around movement onset, or that excitation outstrips inhibition. It was also not the case that interneurons tended to invert their direction / condition preference around movement onset, as might be expected if they switched from inhibiting the planned movement to inhibiting competing movements. Instead, interneurons had a broad distribution of correlations between their plan- and movement-period preferences across cells, with the distribution centered slightly above zero. That is, on average they neither maintained their preferences consistently nor switched their preferences consistently between planning and movement. Surprisingly, pyramidal neurons showed a similarly wide range of correlation values. This wide range of correlations is again inconsistent with the output-gating hypothesis, which predicts that the pattern of pyramidal cell plan-period activity should resemble the movement-period activity (merely at sub-threshold levels). Thus, it does not appear that premotor cortex simply creates a pattern of activity in the excitatory neurons while output cells are kept inhibited until the go cue.

2.4.3 Alternative mechanisms for preventing plan activity from driving movement

Some mechanism would appear to be necessary to prevent plan period activity from causing undesired movements. One possibility is that a small subset of interneurons
might act as gates, while most have some other function. However, we did not see support for this in the form of a subset of cells with a pause-like activity pattern (high plan-period firing rates and low movement-period firing rates). Instead, regarding their tendency to pause during movement, we found a unimodal distribution across neurons (Figure 2.7 bottom row).

A second possibility is that PMd may make feedforward connections primarily on inhibitory neurons in M1, essentially making the rest of M1 the ‘output’ neurons in Figure 2.1. There is physiological evidence that some pyramidal neurons in PMd synapse primarily on inhibitory cells in M1, producing feedforward inhibition (Ghosh and Porter, 1988; Keller, 1993; Tokuno and Nambu, 2000). However, we did not find a clear subset of pyramidal neurons with pause-like activity patterns. The feedforward inhibition hypothesis also makes the concrete prediction that we should find pause-like neurons in M1, which presents a good target for future study.

Another possibility is that gating may occur purely downstream of PMd. Local inhibition in M1 could prevent it from responding during planning, and then be released before movement. Other work has suggested that cancellation of movements, a form of gating, could involve lateral inhibition in M1 (Riehle et al., 2006). However, recent work has shown that in M1, as in PMd, inhibition rises rather than falls around the time of movement onset (Merchant et al., 2008).

An intriguing alternative is that the motor system does not use a straightforward mechanism of inhibitory gating at all. Indeed, a blanket output-gating mechanism might be ill-suited to real-world reaching behavior. One must typically plan one movement while executing another, or while maintaining a posture (e.g., in our task the arm had to be supported against gravity). If a feedforward gate exists, it seems it must therefore be movement-specific. The connectivity of PMd also argues against the presence of a blanket output gate: PMd projects to primary motor cortex (M1) and spinal interneurons but not to lower motor neurons (Dum and Strick, 2002). Because both M1 neurons and spinal interneurons show some plan-period activity (Prut and Fetz, 1999), PMd should be expected to send some signals down the spinal cord and into M1 even during the plan period. Such activity presumably modulates reflexes or otherwise prepares the downstream motor system. Lastly, the mean firing rate across
pyramidal neurons barely changed during the planning period. It is thus possible that planning occurs in such a way that there is no ‘excess’ output to gate, particularly for output units (perhaps including the non-canonical waveform units).

Therefore, instead of a gated oculomotor-like system such as that illustrated in Figure 2.1, we might consider PMd and M1 as forming a dynamical system for controlling the arm (Churchland et al., 2006b; Cisek, 2006b; Fetz, 1992; Scott, 2004; Todorov and Jordan, 2002). Since there are many more neurons than muscles, neural activity would likely be higher dimensional than the muscle activity it controls. If so, many directions in the high-dimensional ‘neural space’ would fall in the ‘null space’ of the muscles: changes in neural activity along those dimensions would not produce changes in muscle activity (Churchland et al., 2007; Yu et al., 2009). An output-gating mechanism would not be required. Such a system would also not be obliged to possess elements with consistent ‘tuning’ during the plan and movement epochs (or even within the movement epoch; Churchland et al., 2006b; Fu et al., 1995; Hatsopoulos et al., 2007). Our findings are consistent with this view, but they fall well short of an explicit test. Future work will have to determine how to generate specific predictions from the dynamical systems / feedback-control framework.

2.4.4 Summary

We found a number of differences between the responses of putative interneurons and putative pyramidal neurons. Two findings present a potential challenge to our current conceptualization of PMd function. First, we found no evidence that interneurons in PMd act to ‘gate’ outputs. It thus remains unclear how, in the absence of movement, plan-period activity can be so prevalent in neurons that also appear actively involved in generating movement. Second, we found that most neurons in PMd have cross-condition modulation that is inconsistent between the plan and movement epochs. It is possible that a dynamical systems / feedback-control framework can account for these findings, although it seems just as possible that the correct theoretical framework has yet to be developed.
CHAPTER 2. LACK OF INHIBITORY GATING IN PMD

2.5 Supplementary Methods

2.5.1 Neural recordings

We made neural recordings using single electrodes in monkey J and an implanted multi-electrode array in monkey H. For monkey J, single-unit recordings were made using tungsten single electrodes with 3-7 MΩ starting impedance (Part # UEWLGSEEN1E, Frederick Haer Company, Bowdoinham, ME) driven by a hydraulic microdrive (David Kopf Instruments, Tujunga, CA), as described previously (Churchland et al., 2006b). Electrode impedances were reduced to ∼1 MΩ after piercing the dura. An effort was made to isolate neurons that were active during the plan period. Neural signals were amplified, filtered, and sorted using a Multichannel Acquisition Processor (MAP; Plexon, Dallas, TX). The signal path was: unity-gain buffer (head stage), 154 Hz 1 pole high-pass filter, 100x amplifier, 8.8 kHz 1 pole low-pass filter, 10x amplifier, 30 Hz 1 pole high-pass filter, digital 400 Hz 2 pole high-pass filter, digital 6 kHz 6 pole low-pass filter (Figure 2.11). Before (or occasionally after) recording each new neuron, waveforms were also collected with the digital filters disabled. For most neurons (137), the recording site depth was estimated. The point at which the electrode entered cortex was determined by listening to the noise characteristics while lowering the electrode. For analyses involving depth, we excluded penetrations where this point was uncertain. 163 neurons were recorded in total, with an average of 14 trials per condition.

For monkey H, signals were recorded from the implanted array using the Cerebus system (Blackrock Microsystems, Salt Lake City, UT) and sorted with the Sahani sorting algorithm (Santhanam et al., 2004) as described previously (Zumsteg et al., 2005). Clustering plots generated by the sorting algorithm assisted in identifying single units by hand; identified multi-units were discarded. During performance of the task, standard filter settings were used (Figure 2.11). Before recordings, two minutes of broadband data were also recorded (0.3 Hz 1 pole high-pass filter, 7.5 kHz 3 pole low-pass filter) for use in waveform classification. A single day of recording (H20041119), which yielded 71 single units, was used.
2.5.2 EMG recordings

EMG activity was recorded from monkey J using hook-wire electrodes (44 gauge with a 27 gauge cannula; Nicolet Biomedical, Madison, WI) placed in the muscle for the duration of single recording sessions. EMG was recorded for a subset of the maze conditions. Recordings were made from trapezius, latissimus dorsi, pectoralis, triceps brachii, medial and lateral aspects of the biceps brachii, and anterior, medial and posterior aspects of the deltoid. Recordings were made one muscle at a time, after completion of neural recording. Electrode voltages were amplified, bandpass filtered (150-500 Hz, four pole, 24 db/octave), sampled at 1000 Hz, and digitized. Off-line, raw traces were differentiated (to remove any remaining baseline), rectified, smoothed with a Gaussian (SD of 15 ms) and averaged.

We verified that plan activity was not producing substantial anticipatory changes in muscle activity. EMG activity was typically unmodulated from the baseline period...
to the plan period, or in rare instances, very weakly modulated during the plan period. This is consistent with our previous verifications in similar tasks with previous animals (Churchland et al., 2006c,b).

2.5.3 Classification of neuron types using single-electrode recordings

To find each neuron’s mean waveform, we took up to 300 waveforms, spline interpolated them to 2.5 $\mu$s precision, then aligned each. Alignment was generally performed at the mid-point of the slope between trough and peak, to minimize jitter for both the trough and peak. For the very broadest-waveform neurons, this yielded poor alignment. Thus, when the downward slope of the trough was $>1.5$ times that of the upward slope of the trough, we instead used the mid-point on the initial downward slope (Figures 2.12, 2.13). We found that this procedure yielded alignment as good or better than that provided by alternate methods based on the trough minimum or trough center of mass. Following alignment, we selected the cleanest waveforms for purposes of obtaining the average waveform (spikes were not discarded for other analyses). To do so, we took the standard deviation (SD) of the waveforms at each time point, then accepted only waveforms where (1) the absolute difference from the mean was $<2$ SDs at every time point, and (2) the absolute mean difference from the mean across time was $<0.4$ SDs. These values were chosen by hand as eliminating most noisy and vertically offset waveforms. We averaged these clean waveforms to get the mean waveform that was analyzed for each neuron. We then separated out by hand neurons with mean waveforms whose shapes were not canonical (e.g., were lacking a trough or a post-trough peak; plotted in Figure 2.10A of Results) and rejected neurons whose mean waveform after-potential was very flat and therefore difficult to measure reliably. Rejections were performed blind to the trough-to-peak duration (TTP) distribution and were performed prior to subsequent analyses. We found a bimodal distribution with a TTP threshold (200 $\mu$s) for separating neuron types, essentially identical to that in Mitchell et al. (2007).
Figure 2.12: Waveform alignment procedure. Top row shows the typical case, bottom row shows the procedure when waveforms have a very shallow post-trough upslope. A randomly-selected subset of waveforms are shown for both examples for visual clarity. Waveforms are first spline-interpolated by 10x then aligned on the deepest point of the trough. Their mean is taken and used to assess whether to align on the downward slope of the trough or the upward slope following the trough. To determine this, we find the midpoint of the downswing and the midpoint of the upswing, and compare their slopes. Typically, the upswing is used. If the downward slope is much steeper (1.5x) than the upward slope, the downward slope is used. This primarily avoids poor alignment of very broad waveforms, which can cause flat peaks in the mean that cannot accurately be quantified. The interpolated waveforms are then aligned on the midpoint of the downswing or upswing as appropriate. Next, we select the waveforms near the mean (see Supplementary Methods) to remove noisy waveforms and the occasional incorrect sort (this occurs more frequently when collecting waveforms than normally, since broader filters are used). Finally, we take the mean of these interpolated, aligned, and down-selected waveforms before measuring the trough-to-peak (TTP) duration. These alignment and down-selection procedures are only used when determining the TTP duration; online sorting is used for all other analyses. Examples shown are from monkey J (P71, P93).
2.5.4 Classification of neuron types using multi-electrode array recordings

We extended the TTP method for classifying neuron types to data collected with the Blackrock multi-electrode arrays. To our knowledge, neuron classes have not previously been distinguished using these arrays. To classify these waveforms, we simulated the filters used with our single-electrode recordings, which we knew could yield good separation of narrow- and broad-spiking neurons. To simulate the single-electrode signal path, we took advantage of the broadband array recordings made before some experiments (Figure 2.11). We re-filtered the original broadband data with digital filters to simulate the (non-digital) single-electrode high-pass filters (thick border in Figure 2.11). Spike sorting was still based on the original filtered waveforms, and the waveforms from broadband were classified by re-filtering them to match the standard array filtering. This produced sorted waveforms with filtering similar to that for monkey J, where we found bimodality.

This process does not yield a distribution of waveforms from the array with two clear peaks, though the distribution does appear to be bimodal (see Results). This blurring is presumably because of electrode tip geometry, lower impedance, filtering properties of the electrode tips, or other unknown properties of the arrays. Use of even broader filters (100 Hz 4 pole high-pass) did not yield cleanly separable peaks either.

We do not know what properties of the arrays are responsible for blurring the
bimodality of the distribution. However, based on simulations using our single-unit data, we found that even relatively small amounts of additional filtering destroyed the two-peaked distribution shape. That exercise also revealed that even with such a blurred distribution it was still possible to properly classify neurons (classification after additional filtering always agreed with that before), so long as one was willing to use a small exclusion zone (30 $\mu$s). We thus applied this zone to the array-recorded TTP data, making our thresholds $<185$ $\mu$s and $>215$ $\mu$s. On the assumption that the filtering properties of the two electrode types (and/or the chronic versus acute recording environment) are unlikely to be exceedingly different, this should be conservative. Importantly, we note that the key differences between neurons classified as narrow-spiking and neurons classified as broad-spiking were very similar between the single-unit dataset and the multi-electrode array dataset. This confirms that the choice of threshold and exclusion zone was reasonable. As a control, we subsequently repeated key analyses using different threshold choices, and found that the main effects were only weakly sensitive to the exact choice of threshold.
Chapter 3

Control of functional connectivity in monkey motor cortex

At the time of this writing, this study was in submission at Neuron as “Control of functional connectivity in monkey motor cortex,” with author list: Kaufman MT, Churchland MM, Ryu SI, and Shenoy KV. MTK collected the data, designed and performed all analyses, and wrote the manuscript.

As noted in the previous chapter, motor cortices exhibit substantial activity while preparing movements, yet the arm remains still. In this chapter, we investigated two models for how functional connectivity between motor cortex and muscles could be controlled. First, we tested whether inhibition within M1 could prevent output during motor preparation, while permitting output during movement. In data from M1 of two monkeys, putative inhibitory interneurons’ firing rates increased during movement, instead of reducing as predicted. With the previous evidence against upstream inhibitory mechanisms, it therefore appears unlikely that motor output is ‘gated’ straightforwardly by cortical inhibition. Second, we tested whether the brain could instead exploit the many-to-few neuron-to-muscle relationship, and avoid movement by remaining in ‘iso-force’ (muscle-constant) neural dimensions. We found that the brain indeed prepared primarily in iso-force dimensions, avoiding ‘muscle-potent’ activity patterns. This mechanism obviates the need for a separate, explicit gate during preparation. We speculate that similar mechanisms could operate elsewhere, to permit
rapid control of functional connectivity between brain areas.

3.1 Introduction

Despite remaining physically connected to one another, brain areas must sometimes causally interact and sometimes decouple. How this functional connectivity is controlled remains an important and open question. One place where such rapid coupling and uncoupling is critical is the motor system. Many neurons in both premotor and primary motor cortex (M1) are active during movement preparation (Riehle and Requin, 1989; Tanji and Evarts, 1976; Weinrich and Wise, 1982). Given that activity in these areas causes movement, we seek to better understand how preparatory activity is prevented from inadvertently driving movement.

When a monkey is cued regarding the path of an upcoming reach but required to withhold the movement until a go cue, preparatory activity is present during this delay in both M1 and dorsal premotor cortex (PMd; Riehle and Requin, 1989; Tanji and Evarts, 1976; Weinrich and Wise, 1982). This activity is tuned for a variety of movement parameters (Churchland et al., 2006b; Godschalk et al., 1985; Hocherman and Wise, 1991; Messier and Kalaska, 2000; Riehle and Requin, 1989), predicts reaction time (Churchland et al., 2006c; Riehle and Requin, 1993), predicts movement variability (Churchland et al., 2006a), and if disrupted delays the movement (Churchland and Shenoy, 2007a). Activity in these areas is also closely related to controlling movement itself. In addition to exhibiting activity during movement (Evarts, 1966; Wise et al., 1986), microstimulation in either area is sufficient to evoke movements (Dum and Strick, 2002; Leyton and Sherrington, 1917; Weinrich and Wise, 1982), and pharmacological reduction of inhibition seems to impair withholding premature movements (Sawaguchi et al., 1996). Given the preponderance of evidence that premotor and M1 activity is involved in both preparing and executing movements, theoretical models have posited a ‘gate’ that can prevent preparatory activity from driving movement (e.g., Benjamin et al., 2010; Bullock and Grossberg, 1988; Cisek, 2006a).

We consider two classes of model that might explain how preparatory activity is prevented from causing movements. The first model is that gating is achieved via the
CHAPTER 3. MOVING & NOT MOVING IN MONKEY MOTOR CORTEX

‘transfer function’ between cortical activity and the muscles. That is, the relationship between cortical and muscle activity might be nonstationary or nonlinear. This class of model encompasses the possibility of a threshold or of gating via time-varying inhibition. Evidence has previously been presented that is inconsistent with a simple threshold model: preparatory activity is tuned very differently from movement-related activity in both M1 and PMd (Churchland et al., 2010a; Crammond and Kalaska, 2000; Kaufman et al., 2010), and higher firing rates do not seem to translate into shorter reaction times (Churchland et al., 2006c). Additionally, outputs from PMd do not appear to be suppressed by inhibition during preparation (Kaufman et al., 2010). It remains possible, however, that inhibition internal to M1 prevents output during preparation. Consistent with this hypothesis, feed-forward projections to inhibitory neurons in M1 have previously been found anatomically (Ghosh and Porter, 1988; Keller and Asanuma, 1993; Tokuno and Nambu, 2000). Evidence against this model in rats has been presented more recently (Isomura et al., 2009), but rats do not have a well-defined PMd-M1 separation and have only weakly tuned interneurons, in contrast to monkey (Merchant et al., 2008). Here we test versions of the ‘feed-forward gating’ model in reaching monkeys by examining the activity patterns of putative interneurons in M1, and provide evidence that these versions of the gating hypothesis are unlikely to be active in preventing premature movement.

Given the evidence against two models of inhibitory gating (in PMd and in M1) and against a simple threshold, we examine a novel class of model for preventing premature movement that does not involve an inhibitory gate. This model, the iso-force space hypothesis, holds that the brain exploits the many-to-few neuron-to-muscle relationship. Because of this relationship, any one pattern of muscle output can in principle be generated by many different patterns of neural firing rates. The brain could thus prepare for movement within this redundant (iso-force) space, and produce no change in muscle output. By examining the population of neurons instead of analyzing individual cells separately, we found that the key prediction made by the iso-force space hypothesis holds in our M1 and PMd data.
CHAPTER 3. MOVING & NOT MOVING IN MONKEY MOTOR CORTEX

3.2 Results

3.2.1 Behavior and recordings

Two adult monkeys performed a variant of the delayed reach task called the ‘maze task,’ described previously (Churchland et al., 2010a; Kaufman et al., 2010). Reaches were either straight or curved to avoid virtual barriers, and most trials had an instructed delay (preparatory) period (Figure 3.1). The maze task provided a delayed-reach paradigm which elicited a wider variety of both neural and muscle responses than a simple center-out reach paradigm, but reach curvature and other such parameters were not directly analyzed here. 108 conditions (mazes) were used for each of two datasets from monkey J, and 27 conditions (mazes) were used for each of two datasets from monkey N. Neural data were collected from M1 and PMd using single electrodes or dual 96-electrode Utah arrays (where noted, J array and N array). The recording locations are shown in Figure 3.2A.

3.2.2 Separation of narrow- and broad-spiking neurons

Previous work has established the trough-to-peak duration (TTP) of the spike waveform as among the more reliable methods for distinguishing putative interneurons from pyramidal neurons (Bartho et al., 2004; Isomura et al., 2009; Mitchell et al., 2007). Specifically, the distribution of TTPs is often bimodal, with the briefer mode thought to correspond with inhibitory interneurons and the second mode corresponding predominately with pyramidal neurons (Krimer et al., 2005). Nonetheless, since we cannot know for certain that every neuron with a narrow waveform is an interneuron, or that every neuron with a broad waveform is not, we refer to them as narrow-spiking neurons and broad-spiking neurons to be explicit that we are not directly measuring whether each neuron is inhibitory or excitatory.

We found a bimodal distribution of TTPs in M1 of both monkeys (Figure 3.2C). In monkey J, 25 M1 neurons were classifiable as narrow-spiking and 36 were classifiable as broad-spiking (Figure 3.2, left); an additional 195 neurons recorded in PMd were used to help clarify the TTP distribution (faded bars in Figure 3.2C, left). In monkey N,
Figure 3.1: Behavior and example neuron and muscle responses. (A) The maze task. One real maze is shown. A timeline of the task appears in the top panel of B. The monkey initially touched a central spot, then a target and (typically) a set of barriers appeared. On some trials, two inaccessible distracter ‘targets’ appeared as well. The target(s) initially jittered slightly in place. The Go cue was indicated by cessation of target jitter, the targets filling in, and the disappearance of the central spot. The monkey then had to make a curved reach around the barriers to touch the accessible target. Trial-averaged biceps EMG (B, middle) and the firing rate for one M1 neuron (B, bottom) are also shown. Times are indicated at the bottom of B; T, target onset; G, go cue; M, movement onset. Flanking traces show SEMs. Shown are maze ID1, neuron J-PM211, and EMG recording J-MB29.

27 M1 neurons were classifiable as narrow-spiking and 44 were classifiable as broad-spiking (Figure 3.2, right); an additional 19 neurons recorded in PMd were used to clarify the TTP distribution (faded bars in Figure 3.2C, right). Some neurons recorded were not classified, either because their waveforms lacked a post-trough peak or they exhibited a flattened post-trough peak that could not be measured reliably. The fraction of our recordings identified as narrow-spiking (41% for monkey J; 38% for monkey N) was greater than physiological (20-30%; Connors and Gutnick, 1990) because an effort was made to preferentially isolate neurons with waveforms that appeared to be narrow. The fraction of neurons classified as narrow- or broad-spiking varied somewhat by depth (Figure 3.2B), but no cell-class analyses showed relevant differences when neurons were analyzed separately based on their depths or anterior-posterior locations.
Figure 3.2: Recording locations and neuron classification. Left column: data for monkey J; right column: data for monkey N. (A) Recording locations. Recording locations for monkey J are superimposed on a photograph of his brain, with major landmarks highlighted with thick black lines. Alignment is estimated. For monkey N, recording sites were registered with a photograph of his brain using measurements taken intra-operatively. Blue dots indicate recordings with broad waveforms, red dots indicate recordings with narrow waveforms. Dots are randomly displaced slightly (0.1 mm) to reveal overlapping recordings. (B) Number of narrow- and broad-spiking neurons recorded by depth, for those recordings where depth could be estimated. (C) Distributions of trough-to-peak durations (TTPs). Bright colors indicate M1 recordings, faded colors indicate additional PMd recordings used to help clarify the TTP distribution and for select analyses where noted. Insets: mean waveforms from all classified M1 recordings.
3.2.3 Testing the feed-forward gating model

Evidence has previously been presented that is inconsistent with inhibition suppressing output from PMd (Kaufman et al., 2010). Here, we consider the feed-forward gating model, as illustrated in Figure 3.3A. In this model PMd or other upstream areas activate inhibitory neurons within M1 during preparation, preventing M1 output. During movement, feed-forward inhibition is released, permitting M1 output. The feed-forward gating model makes specific predictions about the pattern of firing rates that should be observed in both pyramidal neurons and inhibitory interneurons in M1 (Figure 3.3B). Most importantly, M1 interneurons would be expected to have high tonic firing rates during both baseline and preparation, then pause during movement. Pyramidal neurons, in contrast, should on average be more active during movement than during preparation.

Peri-stimulus time histograms (PSTHs) from representative broad-spiking and narrow-spiking neurons are shown in Figure 3.3C. Both neurons exhibit complex, time-varying responses during the movement, and these responses are substantially different for different reach conditions. While the broad-spiking neuron arguably resembles the pattern expected from the model, the narrow-spiking neuron (putative inhibitory interneuron) exhibits a pattern of activity nearly the opposite of what was expected from the model. Instead of having a high baseline rate and then pausing during movement, this neuron has a low baseline firing rate then increases its firing for all conditions during movement.

These trends are representative of the population. Figure 3.4A plots population PSTHs for all narrow-spiking M1 neurons (red) and for all broad-spiking M1 neurons (blue), collapsed across all conditions. For monkey J (left), broad-spiking neurons had very little net modulation, with reductions in firing rate nearly as common as increases in firing rate around the time of the movement (with firing rate changes thus averaging nearly to zero). In contrast, narrow-spiking neurons had a strong net positive modulation, again with most of the deviation from baseline occurring around the time of the movement. Monkey N had much weaker differences between the two neuron classes, but showed a similar general pattern: most net modulation occurred around the time of the movement, and narrow-spiking neurons had a stronger net
Figure 3.3: Feed-forward gating model and example recordings. Heavy outlines in illustrations represent active neurons. In this model (A), most M1 neurons are not active during motor preparation (top panel) because of strong inhibition within M1 during the preparatory epoch. During movement (second panel), this internal inhibition declines and PMd activity drives M1 activity. (B) Responses predicted by the model. Pyramidal cells are expected to be more active during movement than during preparation, and interneurons are expected to have high firing rates during baseline and preparation, then pause during movement. (C) PSTHs of a recorded broad-spiking (putative pyramidal) neuron from M1 (top) and of a narrow-spiking (putative inhibitory interneuron) neuron from M1 (bottom). Both examples were selected to be as representative as possible. Half of the conditions were selected randomly for display, to aid clarity. For display only, PSTHs were smoothed with a 30 ms Gaussian. Neurons J-PM167 and J-PM206.
positive modulation than broad-spiking neurons.

Figure 3.4B plots the mean cross-condition modulation. To compute this value at each time point, for each neuron the range of firing rates across conditions was taken at that time point (independently of other times). This acts as a measure of tuning depth. These plots show that, for both monkeys, narrow- and broad-spiking neurons exhibited similarly large amounts of modulation across conditions (depth of tuning). Together with the population PSTHs, these results indicate that both cell classes are well-tuned, but that narrow-spiking neurons exhibit increases in firing rates more often than decreases, while broad-spiking neurons are modulated positively and negatively about equally (cell-by-cell analyses shown in Figure 3.5). These results are consistent with previous findings in monkey motor cortex (Merchant et al., 2008).

The key prediction of the feed-forward gating hypothesis is that many interneurons should have high firing rates during preparation and reduce their activity during movement. It is clear from Figure 3.4A that this was not the case on average. However, since interneurons are a heterogeneous group (Kawaguchi and Kubota, 1997; Markram et al., 2004), we wished to determine whether some subset of our narrow-spiking neurons might more strongly resemble ‘gate-like’ neurons than the average. To assess how gate-like each neuron was, we took each neuron’s mean preparatory-epoch firing rate (50 to 400 ms after target onset) and subtracted its mean movement-epoch firing rate (50 to 400 ms after movement onset).
CHAPTER 3. MOVING & NOT MOVING IN MONKEY MOTOR CORTEX

A

monkey J

mean PSTH
(mean across all conds)

cross-condition modulation
(pref - least pref)

narrow-spiking

broad-spiking

monkey N

B

cross-condition modulation
(pref - least pref)

C

'gate-like' tendency

# of neurons

preparation - move (spikes/s)
Figure 3.5: Basic cell-by-cell analysis of firing rates. Distributions for monkey J (left) and monkey N (right). This analysis summarizes the PSTHs for each neuron. Top row: for each neuron, the mean plan-period firing rate (FR) is taken and the mean baseline-period FR is subtracted. The histogram over neurons is plotted. Red indicates narrow-spiking neurons, blue indicates broad-spiking neurons. Dots indicate medians. Bottom row: same for mean movement-period FR minus mean baseline-period FR.
firing rate (-100 to +200 ms from movement onset). The histograms of this index are plotted in Figure 3.4C. If neurons tended to pause during movement, this value would be strongly positive. Instead, narrow-spiking neurons trend towards having even more negative indices than broad-spiking neurons do ($p < 0.005$ for monkey J, n.s. for monkey N). These distributions also lack any obvious second mode at a positive value, which would have been expected if a subset of narrow-spiking neurons were acting as feed-forward gates.

The feedforward-gating model predicted that inhibitory interneurons in M1 would have high and perhaps un-tuned firing rates during movement preparation, then exhibit a sharp reduction in firing rates around the time of movement. We found that narrow-spiking neurons instead tended to increase their firing rate around the time of the movement, that they have cross-condition modulation (tuning) at least as strong as that for broad-spiking neurons, and that there does not appear to be a distinct subset of narrow-spiking interneurons that behave consistently with the feed-forward gating hypothesis. While further elaborations of these models are possible (for instance, small numbers of highly specific interneurons which synapse primarily on corticomotor neurons), these data provide evidence against several versions of the inhibitory gating hypothesis. Together with previous evidence against output gating in PMd (Kaufman et al., 2010), we felt these negative results compelled us to examine alternative models for prevention of premature movements as well. We therefore investigated a different class of model suggested previously, the iso-force space hypothesis (Kaufman et al., 2010).

### 3.2.4 The iso-force space hypothesis of movement gating

A subset of neurons in M1 and PMd project down the spinal cord, synapsing on spinal motor interneurons and occasionally even on lower motor neurons (Dum and Strick, 2002). Given this direct projection, the simplest imaginable relationship between neural activity and muscle activity would be a linear one: that is, the activity of the biceps would be some linear combination of neural firing rates, the activity of the deltoid would be a different linear combination of neural firing rates, etc.
Figure 3.6: Simplified iso-force space model. If a muscle receives input from two neurons (blue triangles in A), the muscle will respond to the sum of the two inputs. Thus, if the sum is constant (B, “Iso-force axis”), the muscle does not distinguish between input 1 being high and 2 low, or vice versa. Similarly, when the sum changes, the muscle will respond differently (“Muscle-potent axis”). If neural activity followed the iso-force axis during preparation (C, blue lines), the muscle would not change its output; when the neural activity changes in the muscle-potent direction as well, movement ensues (C, green lines).

To illustrate the iso-force space hypothesis, consider the simplified situation shown in Figure 3.6A. Just two excitatory neurons synapse directly on a muscle, and this muscle produces force proportional to the sum of its two inputs. As long as the sum of the two inputs remains constant, the muscle will produce a constant amount of force. The activity of these two neurons can be represented as a point in a two-dimensional firing rate space, and their pattern of activity over time as a trajectory through this space. In this space, the constant-sum line will therefore form an ‘iso-force’ axis (Figure 3.6B, blue line). Only if the input sum changes will the force output change; we term the direction in which the sum changes the ‘muscle-potent’ axis (Figure 3.6B, green line).

The predicted pattern of activity during a delayed-reach task is illustrated in Figure 3.6C. Activity for three different reaches is shown. Each trial starts at a baseline, which lies (by definition) on the iso-force line. During the delay period, the neurons change their rates, but together they remain on the iso-force line (blue traces). Only
after the Go cue (gray dot) does activity ‘turn’ and move in the muscle-potent direction as well (green traces; a similar turn has been seen previously: Yu et al., 2009). Critically, as long as the preparatory activity of these two neurons moves through the space only along the iso-force axis, no change in the muscle output will occur. No further gate would be needed.

Extending this idea to many neurons (dimensions), the key prediction of the iso-force space hypothesis is that preparatory activity will avoid moving along muscle-potent dimensions. We further hypothesized that if neural activity occupies a relatively low-dimensional space (Smith and Brown, 2003; Yu et al., 2006), and if the relationship between neural activity and muscle activity is even roughly linear (Morrow and Miller, 2003), neural data projected onto muscle-potent dimensions should resemble muscle activity. In this case, it should be possible to identify muscle-potent dimensions by regressing neural activity against associated EMG data. We then expect that the neural state during preparation should move less through the muscle-like (putative muscle-potent) dimensions than through the other (putative iso-force) dimensions.

To test this idea, we employed the procedure in Figure 3.7 (additional details in Methods). For both the neural data and EMG data, the trial-averaged rate over time was taken for each condition and smoothed with a 28 ms Gaussian. We then reduced the dimensionality of the neural data to six dimensions using Principal Component Analysis (PCA), and similarly reduced the dimensionality of the movement-epoch EMG data to three dimensions (Figure 3.7-1). Next, we regressed each dimension of the EMG data against the movement-epoch (but not preparatory) neural data (Figure 3.7-2). Linear regression can be interpreted as finding the direction in the space of the predictors that most closely resembles the pattern of outputs; therefore, each of these three EMG regressions yielded the direction in the neural space which most closely resembled that dimension of EMG activity. We orthonormalized the resulting three regression directions (for mathematical convenience), and took these “muscle-like” dimensions as our putative muscle-potent space (mean $R^2$ value with the EMG data was 0.53). Three dimensions in the neural space were ‘left over;’ and these “muscle-unlike” dimensions were taken as the putative iso-force space (mean $R^2$ value with
the EMG data was 0.11). The neural data were then projected into each of these two spaces (Figure 3.7-3). For these analyses, both M1 and PMd neurons were used (152 total neurons satisfied tuning criteria for monkey J, 82 for monkey N; see Controls for the iso-force space analyses below). Additionally, dual-array recordings were analyzed as separate datasets (73 neurons satisfied tuning criteria for J array; 57 for N array).

The projected neural data obtained via this procedure is simply a rotation of the reduced-dimensional neural data, with some linear combinations of neurons identified as iso-force and other linear combinations identified as muscle-potent. These data can therefore be viewed in PSTH format. Example iso-force and muscle-potent dimensions are plotted in Figure 3.8A. In the muscle-potent example (black, right), there is little deviation from baseline during the preparatory epoch. In contrast, the iso-force example (gray, left) exhibits a strong across-condition spread during preparation, and a deviation from baseline. These examples were selected as being especially clear visually, but are quantitatively average (variance ratio 4.2; see below). We note that the procedure used to isolate the iso-force and muscle-potent spaces were performed over movement epoch data only; thus, patterns observed in preparatory data are not a result of the space partitioning method.

To quantify the difference in preparatory activity between the two spaces, we simply took the variance of all the points (across conditions and time) in each space during preparation and late baseline (-100 to +400 ms from target onset spanning condition and dimension; bars at bottom of Figure 3.8A plots). In order to scale the spaces comparably, these values were normalized by the variance during movement (-200 to +600 ms from movement onset). Intuitively, this metric measures both how much preparatory activity moved away from baseline, and the degree to which different conditions ‘spread out’ in the iso-force and muscle-potent spaces.

In all four datasets, the preparatory variance was much greater in the iso-force space (gray) than in the muscle-potent space (black; Figure 3.8B). Averaging across datasets, the mean ratio of iso-force variance to muscle-potent variance was 4.0. This ratio indicates, of the preparatory activity observed, how much was confined to the iso-force space and how much escaped to the muscle-potent space. We reiterate that this test was performed on only preparatory-epoch data, whereas the muscle-potent...
Chapter 3. Moving & Not Moving in Monkey Motor Cortex

Figure 3.7: Method for testing the iso-force space hypothesis. Traces represent the firing rate of each neuron (one per row) across conditions and times (columns). Each row is effectively one neuron’s PSTH with conditions laid end-to-end. The dimensionality of the neural data across both preparatory (black) and movement (red) epochs was first reduced via PCA, preserving six dimensions (1). The dimensionality of the movement-epoch-only EMG data was also reduced, preserving three dimensions. In (2), “muscle-like” dimensions in the neural space were found. To do so, data was regressed from each of the three reduced-D EMG dimensions onto the movement-epoch (reduced-D) neural data. These regressions did not include preparatory-epoch data. Each of these three regressions yielded a “direction” in reduced-D neural space. These directions were orthonormalized, yielding a 3-D muscle-like space that formed the putative muscle-potent space. The remaining three (“muscle-unlike”) dimensions in neural space were taken as the putative iso-force dimensions. The neural data was then projected separately into the muscle-potent space and into the iso-force space (3). N, number of neurons; T_f, full number of timepoints per condition; C, number of conditions; T_m, number of timepoints per condition during the movement epoch.
and iso-force spaces were identified using only movement-epoch data. Thus, ‘training’ and ‘test’ data were non-overlapping. This analysis therefore serves as the key test of the iso-force space hypothesis: preparatory activity lies preferentially in the iso-force space.

In order to determine statistical significance, our primary concern was whether we might be identifying arbitrary directions in neural space as iso-force, in which case large effects might be obtained ‘accidently’ because of uneven distribution of variance in neural space (perhaps due to outlier neurons or other effects). To address this, a bootstrap was performed in which directions in neural space were chosen randomly as iso-force or muscle-potent, and the size of the measured effect was compared with the distribution of random effect sizes. By this criterion, datasets J, J array, and N array were significant (p = 0.02; p = 0.03; p < 0.001, respectively); dataset N was not significant (p = 0.07; 10,000 resamplings each). We note that this

Figure 3.8 (following page): Testing the iso-force space hypothesis. The iso-force space hypothesis holds that preparatory neural activity should not move through muscle-potent dimensions, in order to avoid causing movement. The neural data can be separated into components that lie in the muscle-potent space and components that lie in the iso-force space (using the method shown in Figure 3.7). After doing this, one dimension of the reduced-D neural data can be plotted like a PSTH (a ‘PCA-PSTH’). One such iso-force dimension is plotted in A (left), along with the corresponding muscle-potent dimension (A, right). As expected under the iso-force space hypothesis, the muscle-potent PCA-PSTH exhibits preparatory ‘activity’ that deviates little from baseline, while the iso-force PCA-PSTH exhibits preparatory activity with substantial tuning and a mean not equal to baseline. This difference can be quantified by taking the variance of baseline and preparatory data points (across conditions and times) in the iso-force space and in the muscle-potent space (B; epoch shown by gray/black rectangles in A). As in the example, the variance in the iso-force space (gray bars) was consistently greater than in the muscle-potent space (black bars). Three of the four datasets were statistically significant via a conservative bootstrap (J; p=0.02; J array: p=0.03; N: p=0.07; N array: p<0.001). The factor by which the variance is greater in the iso-force space than in the muscle-potent space is given in the numbers above the bars; the mean variance ratio with this method was 4.0. In order to gain a better estimate of the size of the effect (and thus of the possible strength of gating via this mechanism), the variance ratio was re-estimated when finding the iso-force space using both movement-epoch and preparatory-epoch data (C). Mean variance ratio using this method was 7.6x. J array is dataset J20090918; N array is dataset N20100812. The example shown in A (dataset JA3D2) has a variance ratio of 4.2x.
CHAPTER 3. MOVING & NOT MOVING IN MONKEY MOTOR CORTEX

A

B

C

Conservative

Best estimate

Fraction of variance

2.6x 4.6x 2.2x 6.8x

6.4x 10.0x 7.1x 6.9x

Muscle-potent

Projection (a.u.)

Project ion (a.u.)

2.6x 4.6x 2.2x 6.8x 6.4x 10.0x 7.1x 6.9x

Conservative Best estimate

Iso-force

Muscle-potent

-400 TARG 400 -300 MOVE 600

-400 TARG 400 -300 MOVE 600

-1 0 1

-1 0 1

Projection (a.u.)

-400 TARG 400 -300 MOVE 600

-1 0 1

Projection (a.u.)

-400 TARG 400 -300 MOVE 600

-1 0 1

Projection (a.u.)
class of bootstrap generates a ‘null’ distribution but not error bars, which is why error bars are not shown in Figure 3.8B.

We also considered how this effect develops over the course of the trial. We expect that different conditions will be similar during baseline, and then fan out through the iso-force space (but not the muscle-potent space) shortly after target onset (as in Figure 3.6C). In order to detect this, instead of taking the variance across a large range of times, the across-condition variance was taken at each time point. This analysis yields a timecourse of the variance in the iso-force space (gray) vs. the muscle-potent space (black; Figure 3.9). Again, in all four datasets, there is greater across-condition variance in the iso-force space than in the muscle-potent space during the preparatory epoch (the variance during movement is necessarily equal because of normalization). Using a similar bootstrap as above, datasets J, N, and N array were significant ($p = 0.007; p = 0.02; p = 0.02$, respectively); dataset J array was not ($p = 0.09$). All datasets were highly significant when resampling conditions (error bars in Figure 3.9).

### 3.2.5 Controls for the iso-force space analyses

We considered a number of ways by which these analyses could have produced a spurious effect. One possibility might be that effect size was driven by outlier responses to a small number of conditions. To address this, conditions were resampled 1000 times with the iso-force and muscle-potent directions held fixed. The standard deviation of the resulting distribution is shown as the error bars in Figure 3.9.

A second possibility is that muscle-potent dimensions could have been dominated by neurons that were only active during the movement, while iso-force dimensions were dominated by neurons that were active only during preparation. This scenario was unlikely a priori, since pyramidal tract neurons exhibit preparatory activity at least as often as non-pyramidal tract neurons (Tanji and Evarts, 1976). Nonetheless, to verify that this simpler explanation was not the primary source of the observed iso-force effect, we ensured that the iso-force and muscle-potent dimensions could not be composed from separate sets of neurons: for the above analyses, only neurons that were strongly tuned during both movement (15 spikes/s for J and N, 8 for J array...
CHAPTER 3. MOVING & NOT MOVING IN MONKEY MOTOR CORTEX

Figure 3.9: Observing the iso-force space over time. Since many neurons exhibit preparatory tuning for the upcoming reach, this would imply that neural activity ‘spreads out’ in the iso-force space across different conditions (as in Figures 3.6C and 3.8A). To observe spreading-out of conditions in the iso-force space over time, we measured the across-condition variance at each time point in the putative iso-force space (gray) and muscle-potent space (black). As predicted, in all four datasets, the across-condition variance during preparation was much greater in the iso-force space than in the muscle-potent space. Flanking traces indicate SEMs computed via resampling of conditions. This difference was also significant via a more conservative bootstrap (see Results) for three of the four datasets (J: p=0.01; N: p=0.02; J array: n.s.; N array: p=0.02).
and N array) and preparation (10 spikes/s for J and N, 5 for J array and N array) were included. These criteria were satisfied by most neurons in the parent data for the J and N datasets, and one third to one half of neurons in the parent data for the J array and N array datasets.

A third possibility might be that there was some special property of using exactly three iso-force dimensions and three muscle-potent dimensions. It is important that the dimensionality be neither too low, else some 'left-over’ neural dimensions could actually be muscle-potent, nor too high, or the later EMG regressions will be largely fitting noise. The dimensionality of three was chosen as accounting for 77-92% of the EMG variance. Nonetheless, we verified that effects were similar with EMG dimensionalities of two, four, and five (always using twice these values for the neural data). In 31 of 32 cases (4 datasets × 4 dimensionalities × 2 analysis methods) the expected-direction effect was observed. In the one remaining case, a small and statistically insignificant opposite-direction effect was obtained; no statistically significant opposite-direction effects were found.

A fourth possibility is that the muscle-potent dimensions could have been dominated by contributions from pyramidal neurons, whereas the iso-force dimensions could have been made up mostly of inhibitory interneurons. We compared the magnitude of the coefficients for muscle-potent and iso-force dimensions from narrow- and broad-spiking neurons, and found that the two spaces were not dominated by either cell type. For two datasets, pyramidal neurons had slightly larger coefficients in the iso-force space, and for the other two datasets they had slightly larger coefficients in the muscle-potent space.

A similar possibility is that the muscle-potent dimensions could have been dominated by contributions from M1 neurons, whereas the iso-force dimensions were dominated by PMd neurons. We therefore split each dataset into anterior and posterior sites, and re-tested all our analyses to verify that they were not somehow simply tapping into rostro-caudal differences in tuning. For all eight half-datasets, we again found that more variance lay in the iso-force space than in the muscle-potent space.

Finally, though the bootstrap naturally incorporated the possibility of creating a spurious effect due to the chosen method of normalization, we explicitly re-tested
without normalization. The effect was somewhat smaller but still present in every dataset (Figure 3.10).

### 3.2.6 Estimating the magnitude of the iso-force effect

The above methods find the muscle-potent and iso-force spaces using only movement-epoch data, so that the effect can be tested on the (non-overlapping) preparatory-epoch data. This choice provides a stringent test of the model’s predictions, since the ‘training’ and ‘test’ data are completely separate. The method is necessarily conservative, however, in that it ignores some of the data when identifying the spaces.

To estimate the effect magnitude, a combination of baseline, preparatory, and movement-epoch data were used to produce better estimates of the muscle-potent and iso-force spaces. These re-measured effect sizes for the four datasets are shown in Figure 3.8C. With the improved estimates, the effect size increased to an average of 7.6 times as much variance in the iso-force space as in the muscle-potent space.

### 3.3 Discussion

It has previously been unclear how the substantial preparatory activity in premotor and primary motor cortex avoids causing movement (Green and Kalaska, 2011). Here, we tested two models. First, we tested the feed-forward gating model, which holds that time-varying inhibition in M1 acts as a gate intervening between cortical activity and the muscles. We found that the properties of M1 putative inhibitory interneurons were not consistent with this model. Second, we tested the iso-force space hypothesis. This model posits that cortical activity is related straightforwardly to muscle activity, but that preparatory activity is structured to remain in the iso-force space, and remain neutral in the muscle-potent directions of neural space. We found evidence that preparatory activity does seem to preferentially occupy the iso-force dimensions of neural space, consistent with the iso-force space hypothesis.
Figure 3.10: Non-normalized iso-force analyses. (A) Same analysis as in Figure 3.8B, but without normalization. (B) Same analysis as in Figure 3.9, but without normalization and showing the variance in each space through the movement epoch as well.
3.3.1 Evidence against the feed-forward gating hypothesis

The feed-forward gating hypothesis illustrated in Figure 3.3A makes a clear prediction: activity of inhibitory neurons should on average be high during the preparatory period, and low during the movement period. Instead, we found that activity of our narrow-spiking neurons (putative interneurons) tended to be highest, not lowest, around the time of the movement. Nor could we find a subset of neurons with a clear pattern of pausing during the movement period. Previous work has found a similar pattern in narrow-spiking neurons in PMd (Kaufman et al., 2010), providing evidence against gating of premotor output as well. In principle, it remains conceivable that feed-forward gating is implemented in the spinal cord, though preparatory activity in the spinal cord has previously been reported to be weak (Fetz et al., 2002; Prut and Fetz, 1999). Even with this caveat, however, inhibitory gating does not seem to explain how it is possible that PMd, which synapses heavily on M1 (Dum and Strick, 2002), can exhibit substantial preparatory activity while M1 exhibits much less (e.g., Riehle and Requin, 1989). While further elaborations of the hypothesis remain possible, such as very small numbers of interneurons providing highly specific inhibition on corticomotor neurons, these data provide evidence against several possible models of inhibitory gating.

3.3.2 The iso-force space hypothesis

It has previously been proposed that the motor system can be viewed as a dynamical system, a machine for producing and controlling movement (Churchland et al., 2006b, 2010a; Cisek, 2006b; Fetz, 1992; Scott, 2008; Shenoy et al., 2011; Todorov and Jordan, 2002). With this viewpoint, the simplest possible hypothesis as to how motor cortex activity could relate to muscle activity would be a linear relationship: muscle activity would be a linear combination of neurons’ firing rates (Rivera-Alvidrez et al., 2010). If neural activity formed a higher-dimensional space than muscle activity, changes in the neural state along some muscle-potent dimensions would be reflected in the muscles, while changes within the ‘extra’ iso-force dimensions would not be reflected in the muscles (Figure 3.6; Churchland and Shenoy, 2007a; Kaufman et al., 2010;
Yu et al., 2009). We would therefore expect that during preparation, in which the monkey’s EMG changes little or (typically) not at all, neural activity would explore the iso-force space but not the muscle-potent space.

We directly tested this prediction. We first correlated neural activity with EMG activity to find the muscle-potent space, then took the ‘extra’ neural dimensions as iso-force. As predicted, we found that preparatory activity for different upcoming reaches spanned a larger expanse of the iso-force space than of the muscle-potent space. Moreover, this effect was large: using only linear techniques, we found that more than seven times as much preparatory variance lay in the iso-force space as in the muscle-potent space. This suggests that the brain may not need a dedicated mechanism like a nonlinear or time-varying ‘gate’ or ‘switch’ for regulating the transfer function between motor cortex and the muscles; instead, it may exploit the relatively higher dimensionality of neural dynamics to ‘get ready’ while keeping the lower-dimensional muscle-potent projection constant (and maintain a nearly linear transfer function).

That such a straightforward analysis can reveal this property is perhaps surprising: neurons are well-known to have nonlinear transfer functions, but our methods could search only for linear relationships between neural activity and muscle activity and only for a planar iso-force space. If the iso-force space were in fact slightly curved, our analyses would likely underestimate the true effect strength.

The present results cannot rule out contributions by other mechanisms, such as nonlinearities in the spinal cord, in preventing preparatory activity from causing movement. However, the magnitude of the effects shown here implies that the iso-force space mechanism is capable of performing most of this function. Nonlinearities are almost certainly present in the system, but given the present results they need not be the primary mechanism for controlling functional connectivity in the motor system.

We also note that the dynamical systems framework, together with the iso-force space hypothesis, could potentially shed further light on the role of preparatory activity in general. In this framework, the neural state must follow a particular trajectory in order to produce the proper time-varying outputs to drive the muscles as desired.
Preparation, then, serves to get the motor system to a state from which it can more readily produce the right trajectory. This model implies that preparatory activity should be best described in terms of upcoming movement activity, which was recently shown (Churchland et al., 2010a).

In addition, the existence of the iso-force space mechanism suggests a new view of the movement-triggering process. Instead of overcoming a threshold or releasing an inhibitory gate, it seems that some ‘Go’ process must pull the neural state out of the iso-force space and into the muscle-potent space. The source and precise nature of this signal remain to be discovered. Finally, we speculate that the iso-force space mechanism could apply elsewhere in the brain as well, to permit rapid changes in the functional connectivity of other brain areas. If these regions (or columns, or some other functional unit) tend to have large ‘iso-output’ spaces, this could help explain how regions that are always physically connected could process information independently but communicate when needed.

### 3.3.3 Summary

We examined putative interneurons and putative pyramidal neurons in M1 to see whether interneurons could be acting as an inhibitory gate for movement. We observed average activity patterns that were opposite to those predicted by the gating hypothesis, and were unable to find a distinct subset of neurons that were consistent with inhibitory gating. Together with similar findings in PMd, this provides evidence that inhibitory gating within cortex is not likely to be a major mechanism for preventing preparatory activity from driving movement. We then tested a second model, hypothesizing that preparatory activity occurs in a plane of neural space that is orthogonal to activity patterns that do influence the muscles. We found that preparatory activity does indeed seem to lie preferentially in the iso-force space, supporting this model.
3.4 Methods

3.4.1 Subjects

Animal protocols were approved by the Stanford University Institutional Animal Care and Use Committee. Subjects were two adult male macaque monkeys (Macaca mulatta) trained to perform a variant of the delayed reach task for juice reward. After initial training, we performed a sterile surgery during which the monkeys were implanted with a head restraint and a standard recording cylinder. The cylinders (Crist Instruments, Hagerstown, MD) were centered over caudal PMd, initially estimated using stereotaxic coordinates (12-15 mm anterior to stereotaxic zero, the intermeatal “ear bar” line) and from previous surgeries and MRIs in other monkeys. The cylinders were placed surface normal to the skull, which was left intact and covered with a thin layer of dental acrylic. To accommodate recording, 3 mm holes were drilled later under ketamine/xylazine anesthesia. After single-electrode recordings were completed, two 96-electrode silicon arrays (Blackrock Microsystems, Salt Lake City, UT) were implanted in M1 and caudal PMd of both monkeys, as estimated from local anatomical landmarks. Array recordings yielded strong, well-tuned preparatory-period activity for arm movements.

3.4.2 Task apparatus

The task apparatus has been described previously (Churchland et al., 2006c). Briefly, during experiments monkeys sat in a customized chair (Crist Instruments) with the head and left arm restrained. Stimuli were back projected onto a frontoparallel screen ~25 cm from the eyes. A photodiode was used to record the timing of video frames with 1 ms resolution. The position of an infrared-reflective bead taped to the fingers was tracked optically in the infrared (Polaris system; Northern Digital, Waterloo, Ontario, Canada). The eyes were also tracked in the infrared (Iscan, Burlington, MA). A tube dispensed juice rewards.
3.4.3 Task design

Both monkeys performed a variant of the center-out delayed-reach task, called the ‘maze’ task, similar to that described previously (Kaufman et al., 2010). Here, the maze task is used simply as a 27- or 108-condition delayed-reach task, but details are given below for completeness.

Experiments consisted of trials, each a few seconds long, which ended in a juice reward if successful. The animal began a trial by fixating (for at least 700 ms) a central fixation spot with his eyes and a cursor projected on the screen a few centimeters above his hand. On 1/3 of trials, a single target appeared. On another 1/3 of trials, a target and up to 9 rectangular barriers appeared. The last 1/3 of trials was identical to the previous type but an additional 2 distracter targets appeared as well. After a preparatory period (0-1000 ms), a go cue was given, and reaches were rewarded if they were swift and did not pass through a barrier. Reward was delivered after the target was held for 450 ms (monkey J) or 700 ms (monkey N), with the next trial beginning a few hundred milliseconds later. When the targets first appeared, they were hollow and jittered slightly (2-3 mm). The go cue was indicated by cessation of target jitter, the targets filling in, and the extinguishing of the fixation spot. A variety of catch trials were also interleaved, including randomly generated novel mazes.

3.4.4 Neural recordings, classification of neuron types, and EMG

Neural and EMG recordings were made using previously described techniques (Churchland et al., 2006c; Kaufman et al., 2010). Briefly, except where noted, neural recordings were made one at a time using tungsten single electrodes (Frederick Haer Company, Bowdoinham, ME) and a Plexon Multichannel Acquisition Processor (Plexon, Dallas, TX). Neurons were classified as narrow- or broad-spiking based on their waveform shapes using previously described methods (Kaufman et al., 2010). In both monkeys, we preferentially isolated neurons that appeared to be narrow-spiking to reduce the total number of recordings required. Recordings were made in the range of medio-lateral
locations that best produced shoulder or upper arm movements when microstimulation was performed. For M1 cell-type analyses 77 neurons were collected for monkey J; 92 were collected for monkey N. These recording locations are shown in Figure 3.2A. For combined M1/PMd analyses (which did not require cell type analysis for all neurons), additional recordings from both areas were used (J: 96 from PMd, 38 from M1; N: 25 from PMd, 1 from M1). For classification of waveforms only, another additional set of PMd recordings was used for monkey J (113 neurons). The median number of trials per maze condition was 13 for monkey J and 11 for monkey N.

While the M1/PMd boundary cannot be identified definitively without histology, monkey J’s M1 recordings appear to be located entirely within 5 mm (anterior) of the central sulcus, which has previously been described as M1 proper (Boudrias et al., 2010). For monkey N, a few recordings were as far as 6 mm anterior of the central sulcus and thus likely in the M1/PMd ‘transition zone’ (Keller, 1993; Weinrich and Wise, 1982; Wise et al., 1986), but these recordings yielded similar results to those posterior when analyzed separately. Microstimulation at multiple sites in both monkeys evoked movements of the shoulder and upper arm, or (much less often) of the wrist.

We also performed recordings using two 96-electrode ‘Utah’ arrays in both monkeys. In both cases, one array was placed in surface M1 and the other over caudal PMd or the PMd/M1 boundary. These datasets are referred to as “J array” and “N array,” by monkey.

EMG recordings were made using hook-wire electrodes (44 gauge with a 27 gauge cannula; Nicolet Biomedical, Madison, WI) inserted into the muscle for single recording sessions, which were interleaved with the neural recording sessions. For monkey J, recordings were made from trapezius, latissimus dorsi, pectoralis, triceps brachii, medial and lateral aspects of the biceps brachii, and anterior, medial and posterior aspects of the deltoid. For monkey N, recordings were made from proximal, middle and distal aspects of the trapezius, latissimus dorsi, pectoralis, triceps brachii, medial and lateral aspects of the biceps, and anterior, medial and posterior aspects of the deltoid.

For both monkeys, the triceps was modulated very weakly during the task and
these recordings were therefore not included in subsequent analyses. For monkey N, the latissimus dorsi was also very weakly modulated during the task and thus was excluded as well.

EMG electrode voltages were amplified, band-pass filtered (150-500 Hz, four pole, 24 db/octave), sampled at 1000 Hz, and digitized. Off-line, raw traces were differentiated (to remove any remaining baseline), rectified, smoothed with a Gaussian (SD of 15 ms) and averaged.

3.4.5 Details of the iso-force space analysis

Analysis was restricted to neurons that were well-tuned during both preparation and movement. The criteria used were 10 spikes/s during preparation and 15 spikes/s during movement for single-electrode recordings, and 5 spikes/s during preparation and 8 spikes/s during movement for the array recordings (tuning assessed as the firing rate range across conditions). For the arrays, spike sorting was performed offline using a custom software package, and only single-unit isolations and clear, stable multi-unit isolations were used. For single-electrode recordings, all isolations were single-unit.

To find the neural ‘state space,’ a matrix (the “A” matrix) was built. Rows of A correspond to neurons, and columns correspond to time and condition. This matrix is thus of size \(N \times T_f \times C\), where \(N\) is the number of neurons, \(C\) is the number of conditions, and \(T_f\) is the ‘full’ number of time points for that condition. Data from the baseline and preparatory epochs (-100 to +400 ms from target onset) and the peri-movement epoch (-200 to +600 ms from movement onset) was used, and each row was normalized by its range. PCA was then run on the neural A matrix to reduce the dimensionality, and only the first six dimensions were retained. The resulting matrix was thus of size \(6 \times T_f \times C\).

To find the ‘muscle-potent’ subspace of this neural state space, we found directions in the neural state space that resembled the EMG activity we recorded. We reduced the dimensionality of the EMG data by producing an EMG A matrix and running PCA as above, retaining three dimensions. Each of these three dimensions was regressed separately onto the neural data, with the muscle activity advanced by
50 ms (Morrow and Miller, 2003). Regularized regression was used to reduce overfitting, with regularization sufficient to reduce $R^2$ values by $\sim 10\%$. The mean norm of the resulting regression coefficients was 0.52, indicating that regression did not overfit the data. We note that the resulting correlation values ($R^2$ of 0.11-0.79, mean 0.53) were somewhat modest; this occurred both because the regression was performed with low-D neural data, and because regularization was used in the regression. These regressions yielded the three (not perfectly orthogonal) directions in the neural space which individually most resembled the EMG activity patterns. This (neural) space was then orthonormalized, and the resulting subspace is referred to as the ‘muscle-potent’ subspace. The remaining three neural dimensions, which corresponded less well to the EMG activity, are referred to as the ‘iso-force’ subspace.

The values of interest were the across-condition variances of the reduced-D neural data within the muscle-potent and the iso-force spaces. For the analyses in Figure 3.8, the non-normalized variance was given by

$$Var = \frac{1}{C} \sum_T \sum_D \sum_C (x_{t,d,c} - \bar{x}_d)^2$$

where $T$ is the number of time points, $D$ is the number of dimensions (three except where noted), $C$ is the number of conditions (27), $x_{t,d,c}$ is the reduced-dimension neural data for time $t$ dimension $d$ and condition $c$, and $\bar{x}_d$ is the mean reduced-dimension neural data across conditions and (relevant) times in dimension $d$.

For the analyses in Figure 3.9, the non-normalized variance at a given time point $t$ was given by

$$Var(t) = \frac{1}{C} \sum_D \sum_C (x_{t,d,c} - \bar{x}_{t,d})^2$$

where $\bar{x}_{t,d}$ is the mean reduced-dimension neural data across conditions at time $t$ in dimension $d$.

To calculate the normalized variances, the mean variance for the space (either iso-force or muscle potent) was calculated over the peri-movement period, and the
preparatory variance was divided by this normalizing value. Without this normalization, variance between the two spaces is not necessarily comparable. Nonetheless, all analyses were also performed without normalization (see Results and Figure 3.10) with similar results. For monkey J, each 108-condition dataset consisted of 4 sub-datasets of different mazes (27 conditions per sub-dataset). Since EMG was performed for each sub-dataset in different sessions, we performed our analyses on these sub-datasets separately and averaged the results to obtain the reported values.
4.1 Summary

The previous chapters have considered how preparatory activity can avoid causing movement. This chapter focuses on how preparatory activity translates into movement, and what the structure of the movement activity looks like.
Most theories of motor cortex have assumed that neural activity represents movement parameters. This view derives from an analogous approach to primary visual cortex, where neural activity represents patterns of light. Yet it remains unclear how well that analogy holds. Single-neuron responses appear very complex, and marked disagreement persists regarding which movement parameters are represented. A better analogy might be with other motor systems, where a common principle is the production of rhythmic, oscillatory neural activity. We found that neural responses in motor cortex of reaching monkeys contain a brief but strong oscillatory component, something quite unexpected for a non-periodic behavior. Oscillation amplitude and phase followed naturally from the preparatory state, suggesting an important mechanistic role for the neural activity that precedes movement. These results demonstrate a remarkable degree of simple structure in the population response. That simple underlying structure explains many of the otherwise confusing features of individual-neuron responses.

4.2 Introduction

Motor and premotor cortex were among the first cortical areas to be studied extensively (Evarts, 1966, 1968; Fritsch and Hitzig, 1870; Leyton and Sherrington, 1917; Penfield and Boldrey, 1937), yet their basic response properties remain poorly understood. This is surprising given that motor cortex is closer to the periphery than is visual cortex: a minimum of two (Kuypers, 1960; Landgren et al., 1962) versus four synapses. This is close enough to yield reliable spike-triggered averages between some neurons and muscles (Fetz and Cheney, 1980; Lemon et al., 1986). Yet it remains controversial whether neural activity relates principally to muscles or to more abstract movement features (Aflalo and Graziano, 2007; Evarts, 1968; Fetz, 1992; Georgopoulos et al., 2007; Hatsopoulos, 2005; Kalaska, 2009; Morrow et al., 2009; Mussa-Ivaldi, 1988; Reina et al., 2001; Sanger, 1994; Schwartz and Moran, 2000; Scott, 2000, 2008; Todorov, 2000). The seminal observation that many neurons exhibit preferred directions during reaches (Georgopoulos et al., 1982) cannot by itself resolve this controversy; many muscles also exhibit preferred directions (Mussa-Ivaldi, 1988).
A number of well-designed experiments (Ajemian et al., 2008; Kakei et al., 1999; Scott and Kalaska, 1997; Sergio et al., 2005) have addressed this ambiguity, yet none has resolved the central debate. In particular, no study has found the ‘reference frame’ in which a stable preferred direction reliably explains neural responses during reaching. In fact, the preferred direction frequently differs among earlier and later times during a given reach (Churchland and Shenoy, 2007b; Hatsopoulos et al., 2007).

At the heart of such debates is the complexity of the neural responses themselves (Churchland and Shenoy, 2007b; Churchland et al., 2010a; Fetz, 1992). Individual-neuron responses in motor and premotor cortex are highly diverse and exhibit a variety of multiphasic patterns (Churchland and Shenoy, 2007b; Churchland et al., 2010a). There exist two classes of explanations. First, responses may reflect a multitude of represented movement parameters:

\[ r_n(t) = f_n(param_1(t), param_2(t), param_3(t) \ldots) \] (4.1)

where \( r_n(t) \) is the firing rate of neuron \( n \) at time \( t \), and \( f_n \) is a tuning function specific for each neuron. An alternative suggestion is that motor cortex acts as a dynamical system that generates and controls movement (Churchland and Shenoy, 2007b; Churchland et al., 2010a; Fetz, 1992; Scott, 2008; Todorov and Jordan, 2002):

\[ \dot{r}(t) = f(r(t)) + u(t) \] (4.2)

where \( r \) is a vector describing the population-level firing rate, \( \dot{r} \) is its derivative with respect to time, \( f \) is an unknown function, and \( u \) is an outside input to the system. Downstream muscle activity, \( m \), would then be a function of the population response:

\[ m(t) = g(Wr(t)) \] (4.3)

where \( g \) is an unknown function, and \( W \) is a matrix mapping the response of many neurons onto a few dimensions of muscle activity. While this mapping is important, the central focus of the dynamical approach regards equation 4.2. Of particular importance is the possibility that neural responses primarily reflect the dynamics captured
by $f$, and display ‘tuning’ for movement parameters only incidentally (Churchland et al., 2006b; Cisek, 2006b). The central challenge is whether it is possible to specify simple forms of dynamics that can explain at least some of the observed response complexity in straightforward terms.

To search for simple classes of plausible dynamics, we took inspiration from a common principal for movement generation across the animal kingdom: the production of rhythmic, oscillatory activity (Georgopoulos and Grillner, 1989; Grillner, 2006; Hogan and Sternad, 2007). Although primate motor cortex is presumably much more flexible than a pattern generator with fixed oscillatory output, it is plausible that some basic principles are shared (Drew et al., 2008). By analogy, the visual systems of flies, frogs and primates are very different, yet all share the basic feature of structured receptive fields. We found that a particular response pattern, population-level firing rate oscillations, was indeed present for multiple motor behaviors. These oscillatory patterns possess a simple structure at the population level, yet are consistent with the complexity of single-neuron responses. These observations further suggest that related dynamic principles may be recognizable across different species and classes of movement.

4.3 Results

4.3.1 Oscillatory responses in different motor systems

There are many examples of neural populations that generate oscillatory responses (Briggman and Kristan, 2008; Brown, 1914; Grillner, 2006; Kiehn, 2006; Marder and Calabrese, 1996; Serrano and Miller, 2006). As one example, the medicinal leech generates rhythmic muscle contractions at $\sim 1.5$ Hz during swimming (Kristan and Calabrese, 1976), and many neurons display firing rate oscillations at that frequency (Figure 4.1a) (Briggman and Kristan, 2006). We observed similar oscillatory structure in cortical responses during walking in the monkey: $\sim 1$ Hz oscillations matching the $\sim 1$ Hz movement of the arm (Figure 4.1b, data transmitted wirelessly during walking on a treadmill). Assuming oscillations are generated by population-level dynamics,
two-dimensional projections of the population response should reveal consistent rotations of the neural state (much as the state of a pendulum rotates in a space defined by its velocity and position). Rotations in state space were indeed found for both the swimming leech (Figure 4.1d; projection of 164 simultaneously recorded neurons) and the walking monkey (Figure 4.1e; projection of 32 simultaneously recorded channels). These observations are not trivial (oscillations could be produced in other ways) but the observed rotational structure is largely expected of a neural dynamical system that generates rhythmic output.

The projections in Figure 4.1d,e were obtained via two steps. The first was the application of principal component analysis (PCA) to the population response. Inconveniently, PCA does not always order dimensions in a manner ideal for observing coherent dynamical structure; a plane containing important dynamics could be defined by PC1 versus PC2, but might just as easily be defined by PC1 versus PC3. We thus employ a novel method that finds an informative plane within the top PCs. To be conservative, this ‘jPCA’ method was applied only to the top six PCs, which contain the six patterns most strongly present in the data. Details regarding jPCA will be described below, but the following point is critical. The six jPCs form an orthonormal basis that spans exactly the same space as the first six PCs. If a given pattern is not present in those PCs, it cannot be present in the jPCs.

The central finding of this study is that quasi-oscillatory neural responses are also present during reaches, a decidedly non-periodic movement. This is illustrated by the average firing rate of an example neuron recorded from motor cortex (Figure 4.1c) and by the corresponding population-level projection (Figure 4.1f). The rotation of the neural state is short-lived (∼1 cycle) but otherwise resembles the rotations seen during rhythmic movement. The key question is whether the patterns in Figure 4.1c,f are idiosyncratic—something that might occur by chance—or whether similar patterns are a prominent feature across neurons and reach types.
Figure 4.1: Oscillation of neural firing rates during three movement types. (a) Response of a neuron in the isolated leech central nervous system during a swimming motor pattern. This neuron was one of a population of 164 neurons recorded simultaneously using voltage-sensitive dye. Data are for a 1700 ms segment drawn from a longer period of swimming. Responses (not averaged across repetitions) were filtered with a 100 ms Gaussian kernel before further analysis. (b) Multi-unit response from one of 96 electrodes implanted in the arm representation of caudal premotor cortex. Data from 32 such channels was wirelessly transmitted during walking. Data are for a 2000 ms segment drawn from a longer period of walking. Responses (not averaged across repetitions) were filtered with a 100 ms Gaussian kernel. (c) Response of one of 118 neurons recorded from motor cortex of a reaching monkey (N) using single-electrode techniques. The firing rate was smoothed with a 24 ms Gaussian and averaged across 9 repetitions of the illustrated leftwards reach. (d) Projection of the leech population response into a 2-dimensional space. The jPCA projection is a two-dimensional projection of the first six PCs. The two projections are plotted versus each other (top) and versus time (bottom). Units are arbitrary but fixed for each axis relative to the other. (e) Similar two-dimensional projection for the walking monkey. (f) Similar two-dimensional projection for the reaching monkey.
4.3.2 Firing rate oscillations are prevalent across neurons

To assess the prevalence of quasi-oscillatory responses during reaching we analyzed recordings from 469 single-neuron isolations from motor and premotor cortex of four monkeys (A,B,J,N). We made a further 146 simultaneous recordings (single and multi-unit isolations) from two implanted 96-electrode arrays (monkey J). Monkeys executed either straight reaches (monkeys A and B) or a mixture of straight and curved reaches (monkeys J and N). An instructed delay paradigm was used, allowing monkeys to prepare their reaches before a go cue was given. We analyzed 8 datasets, each employing a unique combination of 27-108 reach types (‘conditions’). For each neuron we computed and analyzed the average across-trial firing rate for each condition (Churchland et al., 2010a).

Most neurons exhibited preparatory activity during the instructed delay, and nearly all responded during movement (Figure 4.2). Single-neuron responses were typically complex, multiphasic and heterogeneous (Churchland and Shenoy, 2007b). Yet one can discern what appear to be short-lived oscillations in the firing rate. These begin shortly before movement onset and are sustained for \( \sim 1-1.5 \) cycles. These quasi-oscillatory patterns were seen for straight reaches (monkey A,B) and for the mixture of straight and curved reaches (monkey J,N). However, considerable interpretational caution is warranted: even random multi-phasic responses might be mistaken for short-lived oscillations when observed at the level of individual neurons or individual conditions. The critical question is whether there is orderly structure at the population level that is consistent across conditions.

4.3.3 Oscillations behave lawfully across conditions

We have proposed that neural activity during movement reflects the evolution of a dynamical system, and that preparatory activity may provide the initial state of that system (Churchland et al., 2006b; Churchland and Shenoy, 2007b; Churchland et al., 2010a; Shenoy et al., 2011). This proposal makes two predictions. First, for a given class of movement (e.g., reaches) dynamics may be reasonably similar across conditions. In the present case, this implies that the neural state should rotate in the
Figure 4.2: Firing rate versus time for 10 example neurons, chosen to highlight the multiphasic patterns commonly observed just before and during movement. Each trace plots the mean across-trial firing rate for one of 27-108 conditions. Traces are colored from red to green based on the firing rate at the end of the preparatory period. Data were averaged separately locked to target onset, the go cue, and movement onset. To aid viewing, traces have been interpolated across the gaps between these three epochs. Scale bars indicate 20 spikes/s. Insets plot the hand trajectory for each condition.
same direction for all conditions. Second, while rotation direction should be the same for all conditions, rotation phase should not be the same, but should depend upon the preparatory state (Churchland et al., 2010a).

To test these predictions we projected the population response for all conditions into the same jPC plane. This was done for 200 ms of data, beginning just as preparatory activity gives way to movement-related activity. Focusing on this time-period aids visualization and allows PCA (and thus jPCA) the best chance to capture preparatory tuning, but all observations noted here continue through the movement. The resulting projections (Figure 4.3a-d) show three unexpected features. First, the same jPC plane captures state-space rotations for a majority of conditions. Second, the neural state rotates in the same direction across conditions (the overall counter-clockwise direction of rotation is arbitrary). Third, the rotation phase follows naturally from the preparatory state. The observed state-space rotations do not relate directly to reach curvature. Monkeys A and B executed straight reaches, while monkeys J and N executed a mixture of straight reaches, clockwise-curving reaches, and counter-clockwise-curving reaches (Figure 4.2, insets). Yet the direction of neural-state rotation was consistent across conditions. Rotations thus appear to reflect a pattern of dynamics that is consistent across conditions, rather than the pattern of kinematics per se.

The observed rotations involve structure that is relatively simple and evolves in a consistent way across conditions. In particular, if one knows the initial preparatory state then the subsequent states can be predicted reasonably well. Such predictability is very much not observed at the level of individual neurons, where the correlation between preparatory ‘tuning’ and tuning at the time of the movement is on average nearly zero (Churchland et al., 2010a; Kaufman et al., 2010) (Figure 4.4). Yet this low correlation, among other response features, follows from the rotations seen in Figure 4.3. This can be made patent by plotting each projection as a function of time (Figure 4.3e). The resulting response patterns resemble those of individual neurons, a reflection of the fact that they exist within the top principal components. Those responses contain features, including reversals of the preferred direction, that might appear odd but are in fact a straightforward consequence of the underlying rotations.
Figure 4.3: Projections of the neural population response, produced by applying jPCA to the first 6 principal components of the data. (a) Two-dimensional projection for monkey B (74 neurons; 28 conditions all employing straight reaches). Each trace corresponds to one condition and plots the first 200 ms of activity as it evolves away from the initial preparatory state (circles). Traces are colored green to red based on the projection of the preparatory state onto jPC1. (b) Two-dimensional projection for monkey A (64 neurons; 28 conditions employing straight reaches). (c) Two-dimensional projection for monkey N (118 neurons; 27 conditions employing straight and curved reaches). (d) Projection for the array-based dataset from monkey J (146 single- and multi-unit isolations; 108 conditions employing straight and curved reaches). (e) The projection onto each jPCA axis as a function of time (format as in Figure 4.2) for monkey N. Data are the same as in panel c, except more times are shown. Traces are colored red to green based on the level of preparatory activity for that projection. This allows visualization of ‘tuning’ with respect to the reach trajectories (inset) but also means that coloring is different for the two projections.
Finally, under the dynamical systems view, preparatory activity and then produce perimovement activity represent-
tory activity and then produce perimovement activity. Second, perhaps preparatory and peri-
productivity activity code the same thing (e.g., reach direction)
with unrelated preferences (e.g., different preferred directions).

Figure 4. Correlation between Preparatory and Perimovement Tuning. (A-D) Distribution of correlations (measured once per neuron) for the four datasets. Analysis was restricted to neurons robustly tuned during both epochs (Methods). Red dot indicates the distribution mean. (E) Average correlation as a function of when perimovement activity was assessed. Peri-movement activity was measured at a single time-point, after smoothing with a 20 ms Gaussian kernel. Correlations are initially high, as preparatory tuning is being correlated with itself.
In contrast with the multi-cycle oscillations seen in the swimming leech and walking monkey, the oscillations during reaching are short-lived: 1-1.5 cycles (Figure 4.1c). Given this, the observed patterns are more properly referred to as ‘quasi-oscillatory’. Yet despite their brevity, the patterns observed during reaching share the same basic rotational structure as those during swimming and walking. However, the relatively high frequency (up to $\sim 2.5 \text{ Hz}$) of the brief oscillations makes them easy to miss. Even at the single neuron level (Figure 4.2), the briefly oscillatory patterns are difficult to resolve unless trial-counts are high (at least $\sim 8 \text{ trials/condition/neuron}$; the eight datasets averaged 810 trials/neuron, $\sim 10-30 \text{ trials/condition/neuron}$) and unless averages are made across very similar movements precisely aligned on movement onset (methods). The brevity of the oscillations / rotations also raises the potential concern that such structure—even when consistent across conditions—might be observed by accident or might be present for purely trivial reasons. These possibilities are evaluated below.

### 4.3.4 Controls: jPCA isolates rotational structure only when such structure is present

The central result of this study is the presence of the rotational patterns seen in Figure 4.3. Such patterns are striking, and are unexpected given traditional models of tuning in motor and premotor cortex. Yet might rotational patterns appear by accident, as we are applying a method that seeks such structure? Note that the jPCs are simply an orthonormal basis, and do not involve arbitrary linear combinations of responses or any other type of fitting. It seems unlikely that a projection of 6-dimensional PCA space could ‘invent’ rotational structure consistent across 27-108 conditions. This intuition is born out by a variety of ‘shuffle’ controls, explained below, demonstrating that jPCA does not find rotational structure when such structure is not present. We similarly observed, in the walking monkey, that rotations were present only during walking and were not erroneously found when the monkey was stationary.

The effect of each of three shuffle controls is illustrated for one example neuron in Figure 4.5, and for the whole population in Figure 4.6. These controls were designed to
address the concern that the jPCA method might be powerful enough to find state-space rotations for any population response that contains diverse and multiphasic responses. This would be a large concern were one analyzing only a few conditions in a very high-dimensional space. However, our analyses involved 27-108 conditions. Furthermore, the jPCA algorithm was applied only after using traditional PCA to reduce the dimensionality of the data to six. Thus, jPCA can reveal only patterns of activity that are strongly present in the data. This greatly lessens any potential concern.

Still, it is prudent to empirically evaluate the degree to which rotational patterns

Figure 4.5: Shuffle controls for jPCA on an example neuron. Explanations of the three shuffles are presented in the text, and the population results are shown in Figure 4.6.
Figure 4.6: Effect of the shuffle controls illustrated in Figure 4.5. Each panel plots the jPCA projection of the population response for the J-array dataset (108 reach conditions). The top left panel plots the projection for the original, un-shuffled data (same as Figure 4.3). The other panels plot the projections found when jPCA was applied following the three shuffle manipulations. While many individual trajectories remain curved, the overall robustness of the rotational structure, and of its relationship with the initial phase, is largely lost following shuffling. This indicates that the pattern seen in the top-left panel reflects real underlying structure in the population response, rather than the ability of our method to find such structure by chance. These shuffle controls had similar effects across all datasets.
might be found by chance simply because the analyzed responses are diverse and complex. To evaluate this possibility we performed three shuffle controls. These controls preserve the diversity and complexity of responses, but were designed to interfere with any ‘true’ underlying rotations of the neural state. If robust rotational structure survives in the shuffled controls, then this would indicate that rotations can be found by our methods even when not truly present. This would call into question the relevance of what was reported in the main text. If robust rotational structure does not survive, then that would argue that the rotations in the original data do indeed reflect the deep structure of that data, and would not be found by chance for just any population with multiphasic responses.

To understand how these shuffle controls work, it is necessary to first consider what features must be true of the population response in the presence of real rotations of the population state. First, consider two neurons whose multi-phasic responses have a component at the same frequency (say, 2.5 Hz). For 2.5 Hz rotations to be consistent across conditions, it must be the case that, if the first neuron leads the second neuron for a given condition, then the first neuron also leads for all other conditions. It is for this reason that multi-phasic responses are not on their own sufficient to produce consistent rotations at the population level. To observe consistent rotations, response phases must be coordinated among neurons in a fashion that is consistent across conditions. Second, to match our experimental findings, rotations must have phases that agree with the initial preparatory state. This is far from trivial, as it is very possible to have rotations that do not have this feature. Consider two neurons with identical sinusoidal responses but with different latencies. Plotting one neuron versus the other may produce rotations, but unless preparatory activity is somehow coordinated in just the right way, rotation phase will not agree with the preparatory state. We applied three shuffle controls that preserve surface features of the neural response while disrupting (if present) any deeper phase relationship between neurons, or between preparatory and movement activity. All shuffle controls are based on the distinction between preparatory activity (which is left intact) and peri-movement activity (which was shuffled in three different ways). We picked a timepoint 50 ms after the go cue as the dividing point between preparatory and
peri-movement activity. For the first shuffled control, the pattern of peri-movement activity was inverted for half (selected at random) of the conditions. The inversion was performed around the dividing timepoint, such that continuity with preparatory activity was preserved. This procedure was performed separately for each neuron, and thus disrupts any phase relationships that might have been consistent across neurons. Thus, if the data originally contained deep rotational structure, that structure should be largely lost after shuffling.

The second shuffle control was similar to the first, but inverted peri-movement activity pattern for all conditions. This manipulation is not expected to remove all rotational structure (most such structure is merely sign-inverted). However, this manipulation is expected to largely remove any consistent relationship (assuming there is one) between the preparatory state and the phase of subsequent oscillations. Thus, for the time period of interest, this shuffle control is expected to greatly reduce the consistency of any true rotations, and their relationship with the initial state. If such structure is indeed lost following the shuffle, that will be taken as evidence that such structure was in fact present in the original data, and was not found by chance by the method.

The third shuffle control randomly reassigned the peri-movement activity from one condition to the preparatory activity from another. The same reassignment was performed for all neurons. (Again, the beginning of the peri-movement pattern was simply appended to the final firing rate during the preparatory state, such that there was no discontinuity). Like the second shuffle control, the third shuffle control is not expected to remove all rotational structure (many of the peri-movement activity patterns are altered only modestly by this manipulation). However, any true relationship between rotation phase and initial state is expected to be disrupted. Thus, for the time period of interest, this shuffle control is expected to greatly reduce the consistency of any true rotations, and of their relationship with the initial state. If such structure is indeed lost following the shuffle, that will be taken as evidence that such structure was in fact present in the original data, and was not found by chance by the method.

It is also worth noting that the plane that best captures rotations (jPC1 vs. jPC2)
also captures considerable data variance: an average of 28% across the 8 datasets. This is sizeable, given the rather high dimensionality of the data itself (Churchland and Shenoy, 2007b). For comparison, PC2 and PC3 (which capture the second and third most data variance possible) together capture a similar percentage (30%). Thus, the jPCA projection simply captures patterns that were always present in the top PCs, but were difficult to see because they were not axis aligned.

4.3.5 Response heterogeneity is expected given underlying oscillations

The simple model described above is not a neural network. Yet one can simulate the response of units with firing rates that depend on the short-lived oscillations employed by the model. Unless one insists upon segregated populations, most units will reflect various combinations of the fundamental patterns present in the model. We thus simulated firing rates as depending, with random weights for each unit, on the leading and lagging phases of the two oscillation frequencies. ‘Preparatory’ activity was simply the initial state. Simulated units displayed multi-phasic response patterns of remarkable variety. Thus, the generator model naturally predicts that single-neuron responses should be complex and heterogeneous. In particular it predicts the observed lack of relationship between preparatory and movement-period ‘tuning’ (Churchland et al., 2010a), and the frequent reversals regarding which condition garners the largest response (Churchland and Shenoy, 2007b). This was true even though no attempt was made to fit the model to neural responses. (Indeed, this was avoided: there is no guarantee that the recorded neurons principally control the deltoid). Yet despite the response heterogeneity, underlying state-space rotations were readily captured via jPCA. While no condition-by-condition comparison can be made, the overall rotational pattern is broadly similar between simulated and real data. This highlights the degree to which relatively simple underlying patterns can produce difficult-to-interpret responses at the single-neuron level.

To quantify and compare the strength of state-space rotations for real and simulated data, we measured the neural state \( x \) in the two-dimensional jPCA plane for
each condition and time. We then computed the angle between \( x \) and its derivative, \( \dot{x} \). For pure rotations, the angle between \( x \) and \( \dot{x} \) should be \( \pi/2 \). For the velocity-tuned model, the distribution of angles (Figure 4.7a) peaks slightly to the right of zero (jPCA always finds a plane with some net counter-clockwise motion). In contrast, both the generator model and the neural data have peaks near \( \pi/2 \), indicating a strong rotational component.

This was individually true for all eight datasets (small symbols), although the distribution was noticeably wider for the data, indicating a greater complexity of responses than is present in a model with only pure rotations. We also analyzed the distribution of angular velocities across conditions (not shown). These were similar for the neural data and for the generator model (average angular velocity of 2.1 and 2.4 Hz respectively) but were much lower for the velocity-tuned model (0.67 Hz). Further quantification of rotation strength and consistency is presented below.

Figure 4.7 (following page): Summary statistics quantifying the consistency of the rotational dynamics for real and simulated data. (a) Histograms showing distributions of the angle between the \( x \) and \( \dot{x} \) for simulated and neural datasets. For the jPCA-projected data we measured, for each time and condition, the state, \( x \), and its derivative, \( \dot{x} \). We then assessed the angle between these two vectors (illustrated in the inset). Movement of the state directly away from zero results in angles near zero. Pure rotations result in angles near \( \pi/2 \). Note that although each datum (each \( x \) and \( \dot{x} \) pair) can show any angle, the jPCA plane is chosen to maximize counter-clockwise rotations. The mean angle will therefore always be positive. Distributions are shown across all analyzed times and conditions. Dots at top show the distribution peaks for individual datasets. (b) The jPCA plane was produced by finding the skew-symmetric matrix that provides the best fit to \( \dot{x} = M_{\text{skew}}x \). We also found the unconstrained matrix \( M \) that provides the same fit. The performance of \( M \) (assessed via the R\(^2\) of the fit) is guaranteed to be at least as good as that of \( M_{\text{skew}} \) (\( M \) captures general linear dynamics, while \( M_{\text{skew}} \) captures only rotations). Here we assess how well \( M_{\text{skew}} \) performed relative to \( M \). This is done both for the rank 6 matrices that capture dynamics in all 6 analyzed PCs/jPCs (bottom row) and the rank 2 matrices that capture dynamics in the first jPCA plane (top row). Circles plot performance for individual datasets. Squares give overall averages. Asterisks indicate overall performance significantly different (t-test, \( p<0.05 \)) from that of the neural data.
4.4 Discussion

Rotations of the population state are a common feature of the cortical response during reaching. Such rotations follow naturally from the preparatory state and are reasonably consistent, in direction and angular velocity, across the many different reaches that each monkey performed. This strength and consistency of rotational structure is much stronger than expected by chance or given previous models.

The relative simplicity of the rotational patterns contrasts sharply with the apparent complexity of single-neuron responses. Single-neuron responses typically contain multiple phases, with the preferred direction reversing once or more during the movement (Churchland and Shenoy, 2007b). Furthermore, preparatory tuning at the single-neuron level typically has no discernable relationship with the subsequent pattern of movement activity (Churchland et al., 2010a). Yet as suggested previously (Churchland et al., 2010a) and illustrated in Figure 4.3, underlying rotational patterns can produce exactly these features at the single-neuron level. Motor cortex thus presents an interesting case where aspects of the population response are considerably simpler than their reflection at the single-neuron level.

The state-space rotations during reaching are higher-frequency and much briefer than those during locomotion, yet there is certainly some resemblance. That resemblance may be accidental: perhaps rotational structure arises for entirely different reasons in the two contexts. Alternately, the neural dynamics in the two contexts may bear some relation. The production of oscillatory responses is common in a variety of motor contexts (Llinas, 2002; Yuste et al., 2005), including rhythmic movement and locomotion (Georgopoulos and Grillner, 1989; Hogan and Sternad, 2007). Given that the nervous system can generate oscillatory patterns, and given that such patterns provide a good basis set for producing multiphasic muscle activity, a reasonable hypothesis is that non-periodic reaches are generated using neural mechanisms that overlap with those that generate rhythmic movement (Drew et al., 2008; Yakovenko et al., 2011).

We recently reported results suggesting that preparatory activity sets the initial state of a dynamical system, whose subsequent evolution produces movement activity
(Churchland et al., 2010a). On the one hand, such dynamics must be quite sophisticated if they are to solve the difficult problem of non-linear control (Scott, 2004; Todorov and Jordan, 2002). On the other hand, the present results demonstrate that the population response contains at least some strong and simple features that have a straightforward dynamical interpretation. Yet even if one accepts a dynamical systems interpretation it is worth stressing two caveats. First, while the relevant neural dynamics may be strongly reflected in cortex, the identity of the circuitry that creates those dynamics is unclear. Such circuitry may be purely local, or may involve the recurrent circuitry (Middleton and Strick, 2000) that links motor cortex with the spinal cord, with other critical subcortical structures (Thorn et al., 2010), and with the periphery (Herter et al., 2009). Second, although there are always multiple jPCA planes that contain rotations the population response certainly exhibits non-rotational components that live outside these planes. This is especially true during reaching but is also true during the onset and offset of locomotion (data not shown). Non-oscillatory structure is presumably necessary to initiate and terminate movement, and may also relate to non-linear aspects of the dynamics required for stability (Sussillo and Abbott, 2009) or for feedback control (Scott, 2004; Todorov and Jordan, 2002).

Further investigation of the dynamics that generate movement will require not only additional measurements, but also the ability to simultaneously perturb and observe the neural state. It is hoped that theoretical work will produce models that can capture both the rotational and non-rotational aspects of data (including the mechanisms that initiate and terminate rotations). Yet even in advance of a full understanding, our results argue for a critical role for the dynamical systems perspective. Single-neuron responses in motor cortex appear complex and heterogeneous, which has led to great controversy regarding what they ‘code’ or ‘represent’. Yet may of the response features that seem most baffling from a representational perspective appear very natural and straightforward from a dynamical systems perspective. It thus seems increasingly unlikely that motor cortex contains a simple representation of movement in a definable reference frame (Churchland et al., 2010a; Fetz, 1992; Scott, 2008). At the same time it seems increasing likely that it can nevertheless be understood
in relatively straightforward terms: as an engine of movement that employs lawful dynamics.

4.5 Methods

Recordings from the isolated leech central nervous system were made by K. Briggman and W. Kristan and have been described previously (Briggman et al., 2005; Briggman and Kristan, 2006). Recordings from monkey cortex were made using both a delayed reach task (with head restraint) and from an unrestrained monkey walking on a treadmill (Gilja et al., 2010; Miranda et al., 2010). The tasks for the reaching data have been described previously (Churchland et al., 2006c, 2010a; Kaufman et al., 2010). Neural recordings during reaching were made from M1 and PMd.

4.5.1 jPCA

A full description of the novel jPCA method is available in (Churchland et al., Submitted). Briefly, data is preprocessed by normalizing the neuron’s firing rate range and (optionally) subtracting the across-mean condition at each time point. The data is then compiled into a data matrix, $X$, of size $n \times ct$, where $n$ is the number of neurons, $c$ is the number of conditions, and $t$ is the number of time-points. This matrix simply contains the firing rates of every neuron for every condition and every analyzed time. We then used PCA to reduce the dimensionality of $X$ from $n \times ct$ to $k \times ct$. $k = 6$ for all analyses in the main text, which is conservative given the true dimensionality of the data (Churchland and Shenoy, 2007b).

jPCA is a method for finding the projection (onto an orthonormal basis) that best captures any rotational structure that might be present in the data. jPCA is based upon a comparison of the neural state with its derivative. We computed $\dot{X}_{\text{red}}$, of size $6 \times c(t - 1)$ by taking the difference in the state between adjacent timepoints within each row of $X_{\text{red}}$. We then fit using:

$$\dot{X}_{\text{red}} = MX_{\text{red}}$$  (4.4)
\[ \dot{X}_{red} = M_{skew}X_{red} \] (4.5)

Thus, we are attempting to find matrices \( M \) and \( M_{skew} \) that take the state at each time in \( X_{red} \), and transform it into the derivative of the state in \( \dot{X}_{red} \). \( M \) can be found using linear regression. Finding \( M_{skew} \) requires more complex (but still linear) operations. The quality of the above fits was assessed using \( R^2 \) (e.g., Figure 4.7) which captures the ability to linearly predict the data in \( \dot{X}_{red} \) (across all times and conditions) given the data in \( X_{red} \).

\( M_{skew} \) has imaginary eigenvalues, and thus captures rotational dynamics in the data. The strongest, most rapid rotations occur in the plane defined by the complex conjugate eigenvectors associated with the largest two imaginary eigenvalues. These eigenvectors \( (V_1 \text{ and } V_2) \) are complex, but the associated real plane can be found by:

\[ jPC_1 = V_1 + V_2, \quad jPC_2 = i(V_1 - V_2). \]

The jPCA projection is then \( X_{jPCA} = [jPC_1; jPC_2] \times X_{red} \). \( X_{jPCA} \) is thus a \( 2 \times ct \) matrix describing the neural state, projected onto two dimensions, for every time and condition. For a given jPCA plane, the choice of orthogonal vectors \( (jPC_1 \text{ and } jPC_2) \) within that plane is arbitrary. We therefore selected \( jPC_1 \) and \( jPC_2 \) so that any net rotation was counter-clockwise (the same choice was of course used across all conditions for a given dataset) and so that the spread of preparatory states was strongest along \( jPC_1 \). We also computed the proportion of the total data variance captured by the jPCA plane, in a manner exactly analogous to that for PCA.
Chapter 5

Optogenetics in monkey

The contents of this chapter were published as part of “An optogenetic toolbox for primates” in Nature Neuroscience, vol. 14(3), pp. 387-97, 2011, with author list: Diester I, Kaufman MT, Mogri M, Pashaie R, Goo W, Yizhar O, Ramakrishnan C, Deisseroth K, and Shenoy KV. Additional related work is discussed in the article. MTK collected much of the data, designed and implemented the stimulation protocols, designed and implemented the quantitative and video analyses, and edited the manuscript.

Optogenetics is a technique for controlling subpopulations of neurons in the intact brain using light. This technique has the potential to enhance basic systems neuroscience research and to inform the mechanisms and treatment of brain injury and disease. Before launching large-scale primate studies, the method needs to be further characterized and adapted for use in the primate brain. We assessed the safety and efficiency of two viral vector systems (lentivirus and adeno-associated virus), two human promoters (human synapsin (hSyn) and human thymocyte-1 (hThy-1)) and three excitatory and inhibitory mammalian codon-optimized opsins (channelrhodopsin-2, enhanced Natronomonas pharaonis halorhodopsin and the step-function opsin), which we characterized electrophysiologically, histologically and behaviorally in rhesus monkeys (Macaca mulatta). We present a set of optogenetic tools designed for optogenetic experiments in the non-human primate brain.
5.1 Introduction

The previous chapters have described a number of findings regarding the dynamics of neural circuits and the roles that different cell types play in these dynamics. As in most of systems neuroscience, however, these studies were observational; we recorded neural activity and correlated it with behavior. To fully understand the system, one must be able to do a causal manipulation by causing perturbations.

Two methods have traditionally been available to cause such perturbations. The first, pharmacological manipulations, can be targeted to specific cell types and are very powerful (Hikosaka and Wurtz, 1983), but they act on a time scale of minutes, whereas neurons act on a time scale of milliseconds. The second, electrical stimulation, can be used for temporally more precise, but not cell type-specific, manipulations. Electrical stimulation also does not allow highly controlled inhibition and causes electrical interference that hampers the simultaneous electrical recording of neural signals from the same site.

A new method, optogenetics, addresses these challenges by introducing into neurons light-sensitive proteins that regulate the ion conductance of the membrane. These proteins, encoded by microbial opsin genes, are derived from sources such as archaea bacteria and algae. They allow optical excitation (Boyden et al., 2005; Zhang et al., 2006) or inhibition (Boyden et al., 2005; Zhang et al., 2006) of specific neuron types based on their expression or projection patterns. Moreover, optogenetics allows simultaneous artifact-free electrical recording of action potentials (Gradinaru et al., 2007; Zhang et al., 2009, 2010).

Optogenetics has been applied in a multitude of behavioral and electrophysiological studies in rodents (Adamanidis et al., 2007; Airan et al., 2009; Aravanis et al., 2007; Gradinaru et al., 2007, 2009; Tsai et al., 2009), and an initial study in rhesus monkeys has been successful (Han et al., 2009). However, a number of major challenges remain before optogenetic techniques are ready for broad application in nonhuman primate science, including neural prosthetics research (Gilja et al., 2011).

First, it is necessary to characterize the extent, efficiency, tolerance and pattern of opsin expression in nonhuman primate cortex to facilitate scientific interpretation of
results, and to minimize potential risks. Viral vectors, promoters and opsins are the three relevant agents that need to be tested. In addition, the amount of laser power applied to the brain is of central interest as too much power can lead to thermal damage (Cardin et al., 2010; Kim et al., 2007; Lee et al., 2010). Second, the reliability of optogenetic stimulation and its effect on neural activity and behavior need to be tested to aid in the design of future experiments, and to maximize the chances of experimental success.

Here we address these challenges with a panel of optogenetic tools applied in rhesus macaques and tested with single-unit and local field potential electrophysiology. We also compare the effects of optical, electrical and combined optoelectronic stimulation in motor cortex on passive behavior.

5.2 Results

We injected two monkeys at seven different sites with four different constructs (Figure 5.1). These constructs included the membrane channel channelrhodopsin2 (ChR2) (Boyden et al., 2005), which activates neurons when driven with blue light; the chloride pump enhanced Natronomonas pharaonis halorhodopsin (eNpHR2.0) (Gradinaru et al., 2008), which inhibits spiking when driven with yellow or green light; and a stepfunction opsin (SFO), which is a mutated version of channelrhodopsin (hChR2(C128S)) (Berndt et al., 2009) that puts neurons in a state of increased excitability for many seconds after a brief blue light pulse. This last effect can be reversed by a brief pulse of yellow light. Our promoter choices included two human promoters. hSyn has been implicated in the regulation of neurotransmitter release at synapses, particularly at glutamatergic and GABAergic synapses (Bogen et al., 2009). hThy-1 is a gene for a cell-surface protein, and was originally discovered as a thymocyte antigen. It is also present on the axonal processes of neurons (Caroni, 1997). As primate-appropriate viral vectors, we chose adeno-associated virus (AAV) serotype 2 pseudotyped with serotype 5 (here referred to as AAV5). We injected 1 µl of virus each millimeter from the cortical surface to a depth of 6-10 mm (normal to the brain surface), to test infections across all cortical layers and taking into account
Figure 5.1: Schematic overview of preparation. (a) Left, injection device; right, stimulation, recording and in vivo fluorescence detector outline. A standard recording grid guided an injection needle to the desired location. Small quantities (1 µl) of viral vectors carrying the opsin-fluorophore fusion gene were injected at different depths and sites in cortex. Starting 5 weeks after injections we stimulated the injected sites optically and simultaneously recorded electrical neural activity using a combination of an optical fiber and an electrode (optrode) guided by the same grid as used for the injections. (b) Injection sites and viral vectors in monkeys D and B. Monkey D was injected at five different sites with three different constructs. Along a line through motor (AP 16 mm, ML 6 mm and AP 11 mm, ML 6 mm) and somatosensory cortex (AP 7 mm, ML 6 mm) we injected AAV5-hSyn-ChR2-EYFP. More laterally, we injected at two different sites in motor cortex, AAV5-hThy-1-eNpHR2.0-EYFP and AAV5-hThy-1-ChR2-EYFP (AP 11 mm, ML 10 mm and AP 16 mm, ML 10 mm, respectively). Monkey B was injected with AAV5-hThy-1-ChR2-EYFP in somatosensory cortex (AP 6 mm, ML 14 mm) and with Lenti-hSyn-SFO-EYFP in parietal cortex (AP 2 mm, ML 14 mm).
potential cortical folding. Constructs and injection locations are shown in Figure 5.1. All AAV5 vectors had a titer of $10^{12}$ particles per ml. The lentivirus injection in monkey B had a titer of $10^9$-$10^{10}$ particles per ml. Between weeks 5 and 12 after viral vector injection, we optically stimulated the injected sites while simultaneously recording neural activity. We also monitored potential effects of the optical stimulation on passive motor behavior, and compared and combined optical stimulation with electrical stimulation to explore effects on behavior. Five months after injection (monkey D) and four months after injection (monkey B), we assessed expression levels and patterns with standard histological methods.

5.2.1 Optogenetic inhibition

During the period between 5 and 12 weeks after injection of the eNpHR2.0 vector, we illuminated tissue with green (561 nm) or yellow (594 nm) light. Neurons responded with a rapid reduction in firing rate to pulse trains or continuous green light with latencies of 1-3 ms (Figure 5.2a). Power densities ranged from 3 mW/mm$^2$ to 255 mW/mm$^2$, measured at the tip of the 200µm diameter fiber, which produced estimated power densities of 0.34 mW/mm$^2$ to 27 mW/mm$^2$ at the site of electrical recordings (the electrode tip typically led the fiber by 300 µm). For a quantitative analysis of how individual neurons responded to light, we performed a $\chi^2$-test (criterion $P < 0.01$, $\chi^2 = 3.8415$, 1 degree of freedom) comparing baseline activity with activity during illumination. At the eNpHR2.0 expressing site we found that 38% (55/144) of all recorded single units and 22% (7/32) of all multiunits significantly changed their firing rate in response to green light. Typically, responsive neurons decreased their firing rates (Figure 5.2b). Only fifteen cells responded with an increase in firing rate, presumably due to disinhibitory network effects (that is, optical inhibition of a neuron that inhibited the neuron under observation). To investigate further whether the firing rate increase was an indirect network effect or based on the stimulation of axons originating from ChR2-expressing sites we stimulated the eNpHR2.0 expressing site with blue light. We found that 17 units (single and multi-units) responded to blue light with an increase in firing rates (and 5 units with a decrease).
Figure 5.2: Representative example of electrophysiology results and population summary from the eNpHR2.0-expressing site. (a) Raster plots and PSTHs from a light-responsive single unit. The neuron responded with a rapid reduction in firing rate to a train of green laser pulses (20 Hz, 5 ms pulses). The firing rate decreased markedly 2-3 ms after the light pulse (inset). The same neuron showed a complete shutdown of spiking activity during 70-150 Hz pulse trains as well as during continuous illumination with green light (instantaneous power density 27 mW/mm$^2$, energy density 135 $\mu$J/mm$^2$ per pulse, average power density 9.6 mW/mm$^2$ for 1 s complete shutdown caused by 5 ms pulses delivered at 70 Hz; all values refer to the site of electrical recordings). (b) Scatter plot of firing rates of all single units and multi-units recorded at the hThy-1-eNpHR2.0-expressing area during continuous stimulation versus baseline activity. Empty circles mark non-significant responses to light, filled circles show significant responses. The dashed gray line is the unity-slope line, where baseline firing rate and stimulation firing rate are equal.
CHAPTER 5. OPTOGENETICS IN MONKEY

There was a trend toward longer latencies (6-7 ms as opposed to the 2-3 ms latencies at ChR2-expressing sites; see below), which suggested that an indirect network effect was responsible, but short-latency responses also occurred. Simultaneously with single-unit recordings, we measured local field potentials (LFPs). LFP deflections followed stimulation frequencies, and the polarity of LFP deflections caused by green or yellow light was positive with a negative rebound, as expected from the underlying ion flow. Blue light did not cause LFP modulations.

To test whether eNpHR2.0 expression had an effect on neuronal activity we compared baseline firing rates (that is, without optical stimulation) of light-responsive and light-unresponsive single units. Light-responsive neurons did not differ significantly from light-unresponsive neurons in their spontaneous activity (Wilcoxon ranksum test; $P = 0.3$; median light responsive neurons 2.8 Hz, median light-unresponsive neurons 2.2 Hz).

5.2.2 Optogenetic excitation

We illuminated 127 and 53 single units from sites injected with hSyn-ChR2 and hThy-1-ChR2, respectively, with blue (473 nm) light while simultaneously recording from them. To stimulate neurons while still being able to isolate them, we titrated the light intensities for each individual neuron to a level that caused increased spiking without increasing the background activity to a level that would obscure the waveform of the neuron of interest. This resulted in a wide range of applied power densities ranging from 3 mW/mm$^2$ to 255 mW/mm$^2$ at the tip of the fiber (estimated power density of 0.25 mW/mm$^2$ to 20 mW/mm$^2$ at the electrical recording site). Neurons at ChR2-expressing sites responded strongly for all tested frequencies of light pulses (Figure 5.3a,b). Neurons were able to follow 20Hz stimulations (300 $\mu$s to 1 ms pulse width, energy densities of $<0.25$ $\mu$J/mm$^2$ to 20 $\mu$J/mm$^2$) with average latencies of 3 ms. Spike frequencies increased during optical stimulation, whereas spike waveforms remained unaltered. With increasing stimulation frequency, the probability of each light pulse evoking a spike decreased. For higher frequencies (>50 Hz) and continuous stimulation we often observed an initial burst of activity followed by a reduction in
firing rate. In total, 50% (62/127) of all neurons recorded from sites injected with the ChR2 construct under the control of the hSyn promoter and 45% (24/53) from sites injected with ChR2 construct under the control of the hThy-1 promoter responded significantly to blue light. A slightly higher percentage of multi-units passed the significance criterion (hSyn: 54/87 (62%); hThy-1: 14/23 (61%); $\chi^2$-test, criterion $P < 0.01$, $\chi^2 = 3.8415$, 1 degree of freedom). The responses were mainly excitatory, with rare exceptions (three single units and two multiunits from hSyn-ChR2 and two single units and two multiunits from hThy-1-ChR2-expressing sites showed overall suppression of spiking activity for at least one of the tested frequencies; Figure 5.3c,d). Simultaneously measured LFPs revealed opposite polarities to LFPs evoked at the eNpHR2.0 injected site: deflections caused by blue light were negative with a positive rebound at sites expressing ChR2. Again, this is as expected given the ionic currents that result from illumination of ChR2 and eNpHR2.0.

The light-responsive neurons did not differ significantly from light-unresponsive neurons in their spontaneous activity in areas injected with AAV5-hThy1-ChR2-EYFP (Wilcoxon ranksum test, $P = 0.64$; median light-responsive neurons 6.5 Hz, median light-unresponsive neurons 7.9 Hz). In areas injected with AAV5-hSyn-ChR2-EYFP, we found a slight reduction in baseline activity of light-responsive neurons ($P = 0.04$; median light-responsive neurons 1.1 Hz, median light-unresponsive neurons

Figure 5.3 (following page): Representative examples of electrophysiology results and population summary from the ChR2-expressing sites. (a) Raster plots and PSTHs from a light-responsive single unit at a site injected with AAV5-Syn-ChR2-EYFP stimulated with a power density of 0.1 mW/mm$^2$. (b) Raster plots and PSTHs from a light-responsive single unit at a site injected with AAV5-hThy-1-ChR2-EYFP stimulated with a power density of 2.6 mW/mm$^2$. The pulse-triggered average is plotted with 1-ms resolution illustrating the latency of the suppression or excitation after the light pulse (insets in 20 Hz panel). The reported mean number of spikes per light pulse is a measure of the reliability with which spikes were evoked, corrected for spontaneous spike rate, averaged across the whole stimulation period and all trials. Spikes per pulse represent averages across all trials. (c,d) Scatter plots of firing rates of all single units and multi-units recorded at areas injected with hSyn-ChR2 and hThy-1-ChR2, respectively, during continuous stimulation. Firing rates during stimulation are plotted against baseline firing rates (without stimulation). Empty circles mark non-significant responses to light, filled circles significant responses. The dashed gray line is the unity-slope line.
a. AAV5-hSyn-ChR2-EYFP

- 20 Hz (1-ms pulse width)
  - Spikes per pulse: 0.61

- 70 Hz (1-ms pulse width)
  - Spikes per pulse: 0.17

- 150 Hz (1-ms pulse width)
  - Spikes per pulse: 0.14

- Continuous stimulation (200 ms)
  - Spikes per pulse: 1.05

b. AAV5-hThy-1-ChR2-EYFP

- 20 Hz (1-ms pulse width)
  - Spikes per pulse: 1.62

- 70 Hz (1-ms pulse width)
  - Spikes per pulse: 0.56

- 150 Hz (1-ms pulse width)
  - Spikes per pulse: 1.05

- Continuous stimulation (200 ms)
  - Spikes per pulse: 0.56

c. AAV5-hSyn-1-ChR2-EYFP

- Baseline firing rate (spikes per s)

- Firing rate during stimulation (spikes per s)

- Firing rate (spikes per s)

- Time (s)

- Latency (ms)

- Frequency evoked spikes

- Firing rate during stimulation (spikes per s)

- Baseline firing rate (spikes per s)

- Firing rate (spikes per s)

- Time (s)
5.2.3 Effect of optical stimulation on passive movements

It has been reported that optical stimulation of ChR2 in the macaque frontal eye field does not cause overt movements (Han et al., 2009). To determine whether this is true for motor and somatosensory cortex, we monitored the contralateral arm and hand during optical stimulation in both monkeys (Figure 5.4). We analyzed 18 stimulation sessions for monkey D and 8 sessions for monkey B. Despite the strong neuronal responses to light, which we recorded simultaneously at the same site with optrodes, we saw no effect on spontaneous motor behavior (the resting arm and hand did not twitch or move during optical stimulation). This was the case even at sites where standard intracortical electrical stimulation reliably caused arm and hand movements.

As a potentially more sensitive assay, we also tested whether optical stimulation could modulate the effect of electrical stimulation in motor cortex in monkey D. We therefore performed electrical stimulation using current levels just barely above threshold with and without simultaneous optical stimulation. We found no increase (or decrease) in electrically evoked hand deflections with the addition of blue light stimulation at ChR2-injected sites (Figure 5.4a). Similarly, inhibition with yellow light at an eNpHR2.0 site did not result in a decrease (or increase) in movements induced by electrical stimulation (Figure 5.4b). Finally, we attempted to ‘prime’ electrical stimulation trains with preceding optical stimulation. Again, this did not seem to influence the magnitude of hand movements. In summary, despite the fact that stimulation with blue light increased neuronal activity by up to two orders of magnitude relative to baseline activity (Figure 5.3c,d) and that hand deflections were evoked by electrical stimulations at the same site, we did not observe an effect of optical stimulation on passive hand movements. This suggests that there is a mechanistic difference between optical and electrical stimulation. In addition to the hand, we monitored the rest of the monkeys’ bodies during optical stimulation. Stimulation did not result in any reproducible change in body posture or any seizure-like behavior.
Figure 5.4: Lack of effect of optical stimulation on passive behavior. (a) Electrical (50 µA) and optical (250 mW/mm² at the tip of the fiber, 50 Hz, 3 ms pulse width) stimulation were both delivered through the same optrode in premotor cortex injected with AAV5-hThy-1-ChR2-EYFP. The amplitude of hand movement caused by electrical stimulations decreased during the course of trials, as is typical with electrical stimulation (black line). Additional simultaneous (dark gray) or preceding optical stimulation (light gray) with blue light neither increased nor decreased the movement, nor did it prolong the efficacy of the electrical stimulation over the course of trials. Optical stimulation alone did not cause any reliable movement (data not shown; no movement observed despite careful observation). (b) Electrical (35 µA) and optical stimulation (250 mW/mm² at the tip of the fiber, continuous light) delivered to motor cortex injected with AAV5-hThy-1-eNpHR2.0-EYFP. Optical stimulation with green light did not increase or decrease the electrically evoked hand deflections. Note that no graph for ‘optical stimulation only’ is plotted as we did not videotape hand deflection for those trials on days during which we applied the combined optical and electrical stimulation protocol.

5.2.4 Bistable optogenetic excitation

We also recorded activity from 12 single units and 17 multiunits from monkey B at sites injected with LentihSyn-ChR2(C128S)-EYFP. On the basis of results in rodents, we expected to see long-lasting increases in neural activity after brief pulses of blue light. This effect has been described to be reversible by yellow light pulses (Berndt et al., 2009). Owing to the cumulative nature of SFO activity, we further expected that short, low-intensity light pulses would gradually increase the activity of the expressing neurons because of increased depolarization with repeated light pulses.
Four single units (33%) and six multi-units (26%) showed the expected longlasting response to a 10-ms blue light pulse. Although a single 10-ms pulse of blue light (3-255 mW/mm$^2$ at the tip of the fiber) typically evoked 1-3 spikes in light-responsive neurons at ChR2-expressing sites, a single 10-ms pulse of blue light of the same intensity range caused neurons at the SFO-expressing site to increase their firing rates for several seconds (Figure 5.5a-c). On average, 355 spikes were evoked by a single pulse, which makes the SFO more than 100 times as ‘responsive’ as ChR2. A 500-ms illumination with yellow light reversed the effect, resulting in abrupt return of the firing rate to the baseline (Figure 5.5d). Repeated pulsing of blue light (10 ms pulse width, 2 s pulse interval, average power density of 0.1 mW/mm$^2$) caused a stepwise increase in firing rate until it reached a plateau after six light pulses (Figure 5.5e).

5.2.5 Histological evaluation of opsin expression

At the end of the experiment, we used standard histological techniques to determine expression patterns. Both promoters, $h$Thy-1 and $h$Syn, resulted in strong local expression in cell bodies, dendrites and axons as well as in fibers that projected to subcortical structures. We found that 43.9% (1174/2673) of all neurons showed expression within a 3-mm circle centered on the AAV5-hSyn-ChR2-EYFP injection. At the center of injections, both constructs under the control of the $h$Thy-1-promoter showed similar expression levels as AAV5-hSyn-ChR2-EYFP (AAV5-hThy1-ChR2-EYFP: 42.9% or 725/1691 averaged across all layers).

5.3 Discussion

In this study we sought to develop and test optogenetic tools specifically for the needs of researchers using nonhuman primates. In the course of these efforts, we assessed the safety and efficacy of two different viral vectors, two primate-specific promoters and three different opsins on an electrophysiological and histological level. Furthermore, we compared and combined optical and electrical stimulation in motor cortex to evaluate
Figure 5.5: Prolonged activation of spiking with a step function opsin (SFO). (a) Raster plot (upper) and PSTH (lower) of a single unit from 2 s before until 30 s after a single 10-ms blue light pulse (255 mW/mm², estimated power density 20 mW/mm² at the electrode tip, energy density 0.2 mJ/mm²). Five repeated trials are shown. (b,c) PSTHs for all single (b) and multi-units (c) responding to a single 10-ms pulse of blue light. Each color represents one single or multi-unit. Responses are superposed. (d) Neuronal responses of the same neuron as in a. Ten seconds after a single 10-ms blue light pulse, we delivered a 500-ms yellow light pulse (80 mW/mm² at the tip of the fiber, 8.6 mW/mm² at the tip of the electrode, energy density 4.3 mJ/mm²) that reversed the activating effect of the blue light pulse. Five repeated trials are shown. (e) Neuronal responses of the same neuron as in a. A train of ten 10-ms blue light pulses was delivered at 2 s intervals. The firing rate increased in a stepwise manner until it reached a plateau after 6 pulses. Four repeated trials are shown.
its impact on passive movements. With the combination of viral vectors, promoters and opsins reported here and in previous work, reliable optogenetic excitation and inhibition of neural activity in non-human primates seems to be possible. Below we discuss the merits and current limitations of this technique to advance the design of future non-human primate neuroscience and neural prosthetic experiments.

5.3.1 Efficacy of optogenetics in non-human primates

We found that 38-50% of all neurons recorded at AAV5-injected sites were light-responsive, in accordance with the histological finding that 40-43% of neurons expressed the construct. This level of expression might seem low but is comparable with other (non-optogenetic) virus characterization studies (McFarland et al., 2009; Nathanson et al., 2009). There are substantial differences between the different AAV serotypes (McFarland et al., 2009) and virus chimerae promise higher efficacy (Grimm et al., 2008).

Unexpectedly, we did not find cells expressing opsin and EYFP in layers 1 and 2/3 with any of the tested constructs. Expression was almost exclusively restricted to layers 4-6. We found the same reduction of expression in upper cortical layers in sulci, arguing against any surface related explanations (for example, virus escape to the surface). Viral tropism and layer-specific promoter readout alone seem unlikely to be the reason for this expression pattern because AAVs and lentiviruses express in all cortical layers (Nathanson et al., 2009). However, the specific combination of promoter, viral vector and titer might have caused the layer specificity. Although the observed expression pattern is surprising it also offers opportunities, by virtue of allowing deep layer-specific stimulation.

In the horizontal dimension perpendicular to the injection track, expression occurred only in a defined diameter of about 1.5-2 mm around the injection track. However, transport of the opsin to axons can lead to expression many millimeters away from the injected area, thus allowing projection stimulation (Gradinaru et al., 2010). The neuron activation caused by blue light in an eNpHR2.0-expressing site that we encountered was probably caused by such stimulation of axons originating from
ChR2 expressing sites. For nonhuman primate studies, in which suitable promotor fragments for specificity are limited, axonal targeting offers a promotor-independent targeting possibility (Chen et al., 2007; Gordon et al., 2009; Gradinaru et al., 2010).

We demonstrated efficacy and functionality of optogenetic control in the expressing neurons by optical stimulation paired with electrical recordings across months after injections. As expected from the channel kinetics, SFO channels only needed a brief pulse of blue light to be activated for many seconds whereas ChR2-expressing neurons responded with only one spike per light pulse. For ChR2, low-frequency stimulation (<50 Hz) was most reliable in causing action potentials. We sometimes encountered effects opposite to the expected directions (blue light causing inhibition at ChR2-expressing sites and green light causing excitation at eNpHR2.0-expressing sites). This was probably caused by an indirect network effect arising from non-expressing neurons receiving increased or decreased inhibitory input from opsin-expressing neurons. However, the overall network activity, as measured with LFPs, showed a strong effect that was in accordance with the expected ion flow. At eNpHR2.0-expressing sites illuminated with yellow light, Cl\(^-\) is pumped into neurons, causing a relative increase in positive charges (positive deflections of the LFP signal) in the extracellular milieu. At ChR2-expressing sites, Na\(^+\) flows into the cells when blue light is present causing a relative decrease in positive charges (negative deflections of the LFP signal) in the extracellular milieu. This is consistent with population-wide inhibition of neural activity at eNpHR2.0-injected sites and population-wide excitation of neural activity at ChR2-injected sites.

We found that optical stimulation in cortical motor and premotor areas did not evoke movements. This finding contrasts with the large effect of optical stimulation on neuronal activity and the ability to evoke movements by electrical stimulations with an electrode in close proximity to the optical fiber tip. Hence, it would appear that optical stimulation did not perturb the system in some way that electrical stimulation does. It is possible that the observed 40-50% of channel-expressing neurons is not enough to yield a behavioral effect. Employing more effective virus serotypes and chimerae could be a solution (Grimm et al., 2008). However, there are other possibilities. First, the observed specificity for deeper layers contrasts with electrical
stimulation, which is likely to affect all cortical layers. Second, the region of stimulation might have been too small to cause overt movement. Whereas electrical stimulation can affect neurons several millimeters away from the stimulating electrode (Histed et al., 2009) by activating distant non-targeted cells that happen to have axonal projections or collaterals near the electrode, optical stimulation only affects approximately 1 mm$^3$ around the tip of the fiber (Aravanis et al., 2007) owing to the local expression of light-sensitive channels and the limited spread of light. Although this small volume seems to be enough to evoke movements in rodents (Gradinaru et al., 2007), the rhesus monkey brain is approximately 250 times larger and might therefore require a larger volume of activated tissue. The introduction of larger fiber diameters and larger numerical apertures for broader light cones, or multiple optical fibers for multi-site stimulation, could be beneficial. However, this will inevitably cause more cortical damage. Engineered opsin genes designed for enhanced light sensitivity, photocurrent size and red-shifted action spectra (allowing greater light spread in tissue) (Gradinaru et al., 2010) are therefore preferable options to improve the toolbox. Third, it is also conceivable that the frequencies of our optical stimulation were not high enough. Electrical stimulation uses pulse trains at several hundred Hz (for example, 300-350 Hz), and it is possible that specific neuron classes follow these high electrical stimulation frequencies. ChR2-expressing neurons, meanwhile, could only follow lower optical stimulation frequencies reliably, with 20-50 Hz being the maximum for reliably evoking spikes. Future experiments using opsins with faster kinetics such as ChETA (Gunaydin et al., 2010) might allow this possibility to be explored. Fourth, compensation dynamics might have masked the effect of the optical stimulation. Fifth, it has been shown that neural activity in primary motor cortex and premotor cortex can increase without causing movement or even EMG activity (Taylor et al., 2002). Therefore it is conceivable that we did not activate the right neuron populations in the exact way necessary to generate muscle activity. Finally, we focused on passively evoked movements in this study. More sensitive measures of behavior might be required, such as optical stimulation while animals are actively involved in a task. In those settings, electrical subthreshold stimulation (stimulation that does not evoke any overt movement) has been shown to influence behavior
(Churchland and Shenoy, 2007a). A similar protocol might therefore be more likely to reveal an effect of optical stimulation.

The aim of this study was to help to enable safe, reliable and effective experiments using tools designed specifically for nonhuman primates. We believe that the characterization of optogenetics is an ongoing process. However, optogenetic gain- and loss-of-function experiments, such as those described here, might allow a sequence of electrophysiology studies similar to classic pharmacological and electrical microstimulation experiments, with increased temporal resolution and cell-type specificity.

5.4 Methods

5.4.1 Optical stimulations and neural recordings

Monkeys sat in a customized chair with the head restrained. We recorded neural data (Plexon data acquisition system) using single microelectrodes and conventional techniques combined with optical stimulation. We inserted a recording grid (Crist Instrument Co., Inc.), equipped with 1-3 optrodes (200 µm optical fibers; Thor labs) glued to tungsten electrodes (FHC) with the electrode tip leading the fiber by 200-400 µm), fed through blunt guide tubes (cut and smoothed 21G 1 1/2 PrecisionGlide Needle; Becton, Dickinson) and attached to a hand-driven Ruffner microdrive (Crist Instrument), into the recording well. We lowered the optrodes into cortex and recorded without bias (all observed neurons were measured). Optical stimulation was computer controlled using the TEMPO software system (Reflective Computing) driving a Master-8 pulse generator (A.M.P.I.) that in turn control led a blue (473 nm, Sanctity, SVL4730100), green (561 nm, CrystaLasers, CL2000) or yellow (594 nm, Laserglow, LRS594CFF1505) laser. Stimulation frequencies were pseudorandomly interleaved. Filter and amplifier settings were as in Chapter 2. Filter and amplifier settings for LFP recordings were: unity-gain buffer, 3.3 Hz 1-pole high pass filter, 88 Hz 1-pole lowpass filter, 500× amplifier. Neural data were sorted offline with the Offline Sorter (Plexon).
5.4.2 Optical and electrical stimulation for movement analysis

The monkeys sat passively with their heads restrained in a soundproof, dimly lit room. We videotaped the hand movement during stimulations and measured offline the maximal deflection of the hand using frame-by-frame analysis. Only trials in which the monkey did not move before stimulation onset were included. Parameters for electrical stimulations were: 50 $\mu$A, 333 Hz, 78 ms pulse train, bimodal pulses with a 300$\mu$s cathodal pulse followed by a 300$\mu$s anodal pulse separated by 150 $\mu$s. Parameters for simultaneous optical and electrical stimulations were: optical: 984 ms at 50 Hz, 3 ms laser pulse width, 3-255 mW/mm$^2$; electrical: 78 ms at 333 Hz, 50 $\mu$A, starting 920 ms after initial laser pulse. Parameters for optical stimulation preceding electrical stimulations were: optical: 984 ms at 50 Hz, 3 ms laser pulse width, 3-255 mW/mm$^2$; electrical: 78 ms at 333 Hz, 50 $\mu$A, starting 100 ms after the last optical pulse.

5.4.3 Data analysis

Neurons were considered light-responsive if they significantly changed their firing rate during optical stimulation in relation to their baseline activity. The criterion was set at $P < 0.01$, $\chi^2 = 3.8415$, one degree of freedom ($\chi^2$ test) for the post-pulse epoch (see below) versus baseline epoch (period before stimulation) for at least one optical stimulation frequency. In addition, the spike modulation (see below) was required to be larger than 0.1 to rule out tiny effects.

We counted the number of spikes occurring 1-9 ms after the light pulse. This number was averaged across all pulses in the stimulation period and across all trials. We baseline-corrected this value so that a neuron firing at its baseline rate yielded zero and a neuron producing exactly one spike per pulse yielded one.

Response signs were calculated as whether the firing rate during the post-pulse epoch was greater than during the baseline epoch (excitation) or less than during the baseline epoch (suppression).

Pulse-triggered averages were taken using 1-ms bins with the laser pulse times
as reference events. To calculate statistical significance for each cell at each pulse frequency, we performed a $\chi^2$ test with a $2 \times 2$ contingency table. The two factors for the $\chi^2$ were whether each bin contained a spike or not, and whether a bin was in a ‘post-pulse epoch’ or in a baseline epoch. The baseline consisted of the portion of each trial before the first laser pulse. For most frequencies, the post-pulse epoch was the period 1-9 ms after the pulse; when the inter-pulse interval was $<9$ ms, we used from 1 ms to the inter-pulse interval. For continuous stimulation, the post-pulse epoch was the entire stimulation period.

Spike modulation was calculated for pulsed light as the increase or decrease in response during the 1-9ms time window following the light pulse as compared to baseline. Spike modulation is relative to the baseline epoch (no stimulation). If the inter-pulse interval was $<9$ ms (for higher frequencies with shorter inter-pulse intervals, for example, 150 Hz), the time window was shortened accordingly. For continuous light stimulation modulation was calculated as the change in firing rate during the entire stimulation period relative to baseline. The calculation was

$$\text{efficiency} = \begin{cases} \frac{\text{sum(triggered average)}}{\text{length}(\text{post-pulse epoch})} - \text{baseline} \times 1,000 - \text{baseline} \\ \text{if excitatory,} \\ \frac{\text{sum(triggered average)}}{\text{length}(\text{post-pulse epoch})} - \text{baseline} \times \text{baseline} \\ \text{if suppressive,} \end{cases}$$

Latencies were taken as the first bin in the pulse-triggered average more than halfway to the maximum or minimum, as appropriate from the response sign.

### 5.4.4 Population PSTHs

For each neuron, we subtracted the mean baseline activity from the PSTH. These baseline-subtracted PSTHs were then averaged across neurons and plotted ± s.e.m.

### 5.4.5 Waveform analysis

Procedures were as in Chapter 2.
Chapter 6

Preliminary results on motor decision-making

This study was performed in collaboration with Mark Churchland, Stephen Ryu, and Krishna Shenoy, with consultation by Roozbeh Kiani. MTK collected the data, designed and performed the analyses, and wrote the chapter. Continuing analysis is in progress.

6.1 Introduction

The previous chapters have largely focused on mean neural dynamics during preparation and movement. In this final chapter, neural dynamics on single trials are considered. In addition, this chapter focuses on the more cognitively complex task of decision-making rather than single-target instructed reaching.

The world often does not simply instruct what movement to make; frequently, a movement must be selected from among multiple options. Certain classes of decision-making problems have been extensively studied in monkey (for reviews, see Gold and Shadlen, 2007; Schall, 2001; Sugrue et al., 2005) and human (Heekeren et al., 2008). However, essentially all of these experiments have focused on either perceptual or value-based decision-making. Perceptual decision-making experiments have centered around finding a signal amidst imposed or internal noise, such as in detecting the
direction of random dot motion (Britten et al., 1996; Horwitz and Newsome, 1999, 2001; Kim and Shadlen, 1999; Shadlen and Newsome, 1996, 2001), matching the frequency of a vibration on the fingertip (Brody et al., 2003; Romo et al., 2002, 1999), categorizing a noisy image (e.g., Afraz et al., 2006), or detecting the dominant member of mixed odorants (Kepecs et al., 2006; Uchida and Mainen, 2003; Uchida et al., 2006). Value-based tasks have generally involved optimizing choices when one of the choices provides better rewards, in the presence of limited sampling or when the relative values may be changing (Sugrue et al., 2004). As summarized recently (Gold and Shadlen, 2007), in these tasks the monkey’s job is fundamentally to estimate “the (unknown) state of the world, given the noisy data provided by the sensory systems.”

A number of models have been designed to account for the observed behavior in these tasks, including parameters such as choice probability and reaction time (Palmer et al., 2005), weighting evidence from some time points in preference to other evidence, and the existence of changes of mind (Resulaj et al., 2009). A variety of decision-making mechanisms have been proposed, including integration to a decision boundary (e.g., Laming, 1968; Palmer et al., 2005), attractor dynamics (e.g., Wong and Wang, 2006), race to threshold (Smith and Vickers, 1988), and probabilistic population coding (Beck et al., 2008). These models can all account for both the observed behavior and the mean neural data.

While some recent progress has been made separating these models (e.g., Churchland et al., 2011), ideally one would simply observe the state of the brain over time on a trial-by-trial basis (Brown et al., 2004; Deadwyler and Hampson, 1997; Nicolelis et al., 1997). In addition, a number of other classes of question may be answerable with a single-trial view, such as determining whether monkeys spontaneously vacillate or at what level of processing an error occurred. Preliminary insights into both of these questions have been described; vacillation during integration (Horwitz and Newsome, 2001) and evidence for pre-existing biases on error trials (Cisek and Kalaska, 2005; de Lafuente and Romo, 2005; Horwitz and Newsome, 2001) have been observed in single neurons. As discussed in Chapter 1, the possibility of single-trial views has been opened by the newly available combination of techniques for obtaining many simultaneous recordings and for analyzing the data using dimensionality reduction algorithms.
specifically designed for use with neural data (Briggman et al., 2006; Brown et al., 2004; Churchland et al., 2007; Yu et al., 2009). When these techniques are further combined with a more cognitively sophisticated behavioral task, it becomes possible to observe the moment-by-moment state of the brain as a monkey makes decisions.

Here, we chose to examine motor decision-making using recordings from the motor system. It is somewhat unclear whether to expect primary motor (M1) and dorsal premotor cortex (PMd) to be involved in the decision-making process itself or whether they should merely reflect the outcome. However, it is known that premotor cortex is influenced by the presence of multiple options (Cisek and Kalaska, 2002, 2005), that ongoing decision-making processes influence oculomotor structures (Gold and Shadlen, 2000, 2003), and that some perceptual information is fed continuously to motor structures (Spikey et al., 2005). Thus, there is good reason to expect motor and premotor cortex to be influenced by the moment-by-moment decision process.

The task used here, the ‘decision-maze’ task, focuses on a somewhat different regime of decision-making than is typical. First, in many trials, all the information about the two possible options is presented immediately to the monkey. Second, the attractiveness of the two options is primarily determined by their motoric difficulty; the monkey is often able to succeed with either option and will be rewarded equally, but one option is often more physically demanding to obtain than the other. Thus, cost is manipulated more than reward. Finally, a sudden change occurs in some trials, strongly encouraging the monkey to change his mind.

In this preliminary report, we demonstrate appropriate behavior in this novel task, good decode of single trials at a fine timescale, the presence of induced changes of mind, and the existence of spontaneous vacillation.

6.2 Results

6.2.1 Behavior

The decision-maze task used was a novel variant of the delayed-reach paradigm, closely related to the maze task discussed in previous chapters. The monkey initiated a trial
by fixating and bringing a cursor into a central fixation box, then two targets and four rectangular virtual barriers were displayed. Initially, the targets jittered slightly in place, indicating that the monkey was not yet allowed to move. ‘Go’ was indicated by cessation of target jitter, the targets filling in, and the fixation point extinguishing. The monkey was then required to make a brisk movement ending at either target and curving to avoid the barriers. Two different basic mazes were randomly interleaved, but each of these two mazes could assume nine different versions. Specifically, two of the barriers could vary slightly to make each of the two targets easy to reach, difficult to reach, or inaccessible (“impossible”). Four examples, including the movement path for that trial, are shown in Figure 6.1. In addition to the $2 \times 9$ maze versions, on $\sim58\%$ of trials, one barrier could change from any of the three difficulty levels to any other difficulty level at a random latency in the trial (any time during the delay period, reaction time interval, or during the movement).

Choice data from three days in two monkeys are shown in Table 6.1. Datasets are divided up by monkey, day, and base maze. The first four columns of this table demonstrate that the monkeys behaved stochastically, not always choosing the same option for the same maze version. This was true even though there was no incentive to divide their choices up between the two options—either target yielded the same reward. Choices were typically relatively evenly divided between the two targets when the difficulties of the two sides were balanced (easy|easy or diff|diff) and differences between the difficulties of the two options altered his proportion of choices in the expected manner.

Barrier changes occurring in the middle of a trial altered choice probabilities as well. For instance, if the left option changed from difficult to easy, or the right option changed from easy to difficult, it was expected that the monkey would be more likely to choose left than before the change. Assessing the choice probabilities in change trials relative to the choice probabilities in no-change trials yielded the values in the last two columns of Table 6.1. Only trials which both started and ended with both targets accessible were considered in this analysis. In all of these datasets, barrier changes produced the expected changes in choice probability.
6.2.2 Neural decode

On average, neurons’ firing rates distinguished between preparing for the two choices (Figure 6.2). However, in order to obtain a single-trial view of the neural state, spiking noise must be overcome. To denoise the data, one can take advantage of the fact that neurons’ firing rates are correlated. Dimensionality reduction techniques, from the field of machine learning, are ideally suited to this task. Here, we use a variant of Factor Analysis, which treats each recorded spike rate as a noisy measurement of a linear combination of the underlying state variables (factors). The variant we use here, Gaussian Process Factor Analysis (GPFA; Yu et al. 2009), is designed to apply appropriate smoothing across time as well.
CHAPTER 6. PRELIMINARY RESULTS ON MOTOR DECISION-MAKING

Table 6.1: Behavioral choice probabilities. Each row represents a dataset, with J1-2 representing data from monkey J on day one, with the second base maze. In the first four columns, the difficulty of the two sides are represented as difficulty left | difficulty right. In the last two columns, the change in choice probability for cases in which a barrier change occurred is given. Only trials that began and ended with both choices accessible were considered. A decrease in difficulty on the left side or increase in difficulty on the right side were both taken as a leftward-biasing change. Only trials with a change occurring at least 100 ms after target onset and before the end of the delay period were included in the last two columns.

<table>
<thead>
<tr>
<th></th>
<th>p(left choice) easy</th>
<th>p(left choice) diff</th>
<th>p(left choice) diff</th>
<th>p(left choice) easy</th>
<th>Δp for left-biasing change (should be &gt;0)</th>
<th>Δp for right-biasing change (should be &gt;0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1-1</td>
<td>0.48</td>
<td>0.21</td>
<td>0.84</td>
<td>0.06</td>
<td>0.29</td>
<td>0.38</td>
</tr>
<tr>
<td>J1-2</td>
<td>0.32</td>
<td>0.35</td>
<td>0.79</td>
<td>0.12</td>
<td>0.31</td>
<td>0.18</td>
</tr>
<tr>
<td>J2-1</td>
<td>0.68</td>
<td>0.39</td>
<td>0.95</td>
<td>0.01</td>
<td>0.18</td>
<td>0.44</td>
</tr>
<tr>
<td>J2-2</td>
<td>0.43</td>
<td>0.69</td>
<td>0.99</td>
<td>0.04</td>
<td>0.45</td>
<td>0.39</td>
</tr>
<tr>
<td>N-1</td>
<td>0.92</td>
<td>0.87</td>
<td>0.99</td>
<td>0.16</td>
<td>0.37</td>
<td>0.37</td>
</tr>
<tr>
<td>N-2</td>
<td>0.89</td>
<td>0.76</td>
<td>0.99</td>
<td>0.14</td>
<td>0.44</td>
<td>0.51</td>
</tr>
</tbody>
</table>

In this preliminary work, our primary interest has been in determining the state of the monkey’s motor plan at each point in time. That is, we wish to decode his neural data at each timepoint as indicating preparation for a left or right movement. We chose to use a simple decode technique, which has generally yielded adequate performance while remaining easily understandable. This technique, which we term “axis decode,” is illustrated in Figure 6.3a. At each timepoint, we take the reduced-D neural state of each forced-left and forced-right trial recorded that day (faded points in Figure 6.3a). These provide cases where we can be reasonably sure what movement the monkey has prepared. These single-trial states are then averaged for each of the two cases (saturated-color points in Figure 6.3a). The axis connecting these two means is then found. Finally, the single-trial states are projected onto this axis, taking the midpoint between the means as zero. This yields a one-dimensional decode value for each timepoint of each trial, negative for leftward plans and positive for rightward plans.

To verify that this method yields reasonable results, leave-one-out cross-validation
was performed. That is, one forced trial was removed from the dataset, and the decode axis was found on the remaining trials. The decode for the held-out trial was then found using this decode axis. This process was repeated for each trial. The decode over time is shown for these cross-validations in Figure 6.3b. Colors are as in a. From this plot, it is clear that the axis decode method worked with high reliability for this dataset. To quantify performance, Figure 6.3c plots a histogram of the decode values for the last point of each trace, which was centered 90 ms after the Go cue. For this representative dataset, decode performance was 97.5% correct.

Table 6.2 quantifies decode performance for all datasets. As a basic test, the first column shows the same cross-validation statistic as in the example above. In 5 of 6 datasets, decode was reliable; in the final dataset (N-2) decode was poor, and this dataset was therefore excluded from further analysis. This indicates that the axis decode method generally works well (though not perfectly) when performed on test trials similar to the training trials.

More interesting is the extent to which free choices can be predicted based on preparatory activity. The second column of Table 6.2 shows the percent correct decode of free choice trials when the decode axis was determined using forced choice trials. In every case, these numbers are somewhat lower than for the cross-validation of forced choice trials.
Figure 6.3: Decode method and example. (a) Decode was based on the forced-left and forced-right trials. In this cartoon, the reduced-D neural state for each trial is shown as a point in state space, with individual trials in faded colors and means in saturated colors. The axis connecting the means was found, then trials were projected onto this axis. The midpoint between the means was taken as zero. A different axis was permitted at each time point, assessed as time from target onset. (b) Decode over time. Each trace represents a trial, with the monkey’s choice indicated by color (blue, left choice; red, right choice). These traces were found using leave-one-out cross-validation over the forced-left and forced-right trials. Traces end 90 ms after the Go cue. (c) A histogram of the endpoints of the traces in b. In this representative dataset, classification was correct on 97.5% of trials. Dataset J2-1.

There are at least three classes of possible reasons for why this decode performance is lower. First, training and testing were performed on different types of trials; it could be that the decode axis found on forced choice trials was not optimally fit to the free choice data. Second, free choice trials might exhibit greater variability, or a slightly different covariance structure. Third, even a perfect decode of motor cortical activity might not perfectly predict the monkey’s ultimate choice. This would be expected if, for instance, the monkey sometimes changed his mind between the Go cue and movement onset, or could somehow execute a different movement than the movement prepared in the context of a decision task.

Hypothesis 1, imperfect decode axis fit, is easily tested. Column 3 of Table 6.2 shows leave-one-out cross-validation for free choice trials. In most cases, decode performance is better than when the decode axis was found using forced choice trials. This indicates that part of the reason for the weaker decode on free choice trials was imperfect fit of the decode axis for this different type of trial. Nevertheless, performance
CHAPTER 6. PRELIMINARY RESULTS ON MOTOR DECISION-MAKING

<table>
<thead>
<tr>
<th></th>
<th>Leave-one-out forced</th>
<th>Free choices, trained on forced data</th>
<th>Leave-one-out free</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1-1</td>
<td>98.1%</td>
<td>82.4%</td>
<td>90.2%</td>
</tr>
<tr>
<td>J1-2</td>
<td>100%</td>
<td>96.8%</td>
<td>95.2%</td>
</tr>
<tr>
<td>J2-1</td>
<td>97.5%</td>
<td>93.5%</td>
<td>93.5%</td>
</tr>
<tr>
<td>J2-2</td>
<td>100%</td>
<td>97.1%</td>
<td>98.6%</td>
</tr>
<tr>
<td>N-1</td>
<td>87.9%</td>
<td>84.6%</td>
<td>84.6%</td>
</tr>
<tr>
<td>N-2</td>
<td>62.5%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>56.7%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>55.0%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Average</td>
<td>96.7%</td>
<td>90.9%</td>
<td>92.4%</td>
</tr>
</tbody>
</table>

Table 6.2: Axis method decode accuracy at 100 ms after Go cue. <sup>a</sup>Dataset N-2 suffered poor decode performance and was thus excluded from further analysis. This dataset is therefore not included in the average.

remains below that for forced choice trials alone. This argues that imperfect decode fit is unlikely to be the sole reason for the lower decode accuracy on free choice trials. Additional analyses will be required to determine whether the remaining difference is caused by late changes of mind or a different reason.

6.2.3 Examining single-trial decision processes

Some types of trials are especially well-suited to single-trial analysis. We chose here to initially focus on changes of mind during the preparatory epoch, which are not directly visible in the movement. Based on behavior alone, it is possible to infer statistically that this process must occur on some trials, but it is not obvious how to determine which trials or what the time-course of this process is. Nonetheless, the change in choice probability with barrier changes implies that the monkey is very likely to have changed his mind on at least some trials.

One clear case of change of mind can occur when a target changes mid-trial from inaccessible to accessible. For instance, if a trial initially appears as a forced left trial, but the right barrier then changes to make it accessible as well, we may infer that the monkey probably changed his mind if he ultimately chooses to reach to the rightward target. Exactly such an ‘encouraged switch’ trial is illustrated in Figure 6.4a (left
The key trial is shown in thick red and the forced left and forced right trials are shown in fainter colors for context. This trial began as a forced left, but 185 ms later (black dot) the right option became ‘easy’ (as shown in the inset). The monkey apparently began his change of mind approximately 150 ms later. A similar change of mind (in the opposite direction, and later in the trial) is shown in the right panel of Figure 6.4a. While such trials are uncommon, such unambiguous examples do appear.

It is rather less clear, however, whether we should expect the monkey to ever change his mind in the absence of any external change. In the free choice case, it has not previously been possible to determine whether spontaneous vacillation occurs. Figure 6.4b shows the decode over time for two free choice trials in which such a vacillation appears to be unambiguous. In the left panel, the decode is initially well within the normal range for a rightward movement, but ~300 ms into the trial switches somewhat gradually to a leftward plan. In the right panel, the decode appears to be a somewhat weak leftward plan, then switches to a rightward plan. By visual inspection of the decodes for free choice trials with long delay periods (≥400 ms), approximately 6-13% of such trials exhibited clear vacillation or (occasionally) a decode ending unambiguously in the non-selected target.

### 6.3 Discussion

Though it has long been clear that decision-making processes would best be examined trial by trial, it has not previously been possible to do so. Here we combined a novel motoric decision-making task with many simultaneous neural recordings and dimensionality reduction methods to gain such a single-trial view. First, the ability to dynamically bias behavior in this new task was demonstrated. Second, analysis indicated that trials were indeed rather heterogeneous, with the neural trajectory on a fraction of trials quite different from the trajectories on others. Finally, evidence was presented for the existence of spontaneous vacillation (covert changes-of-mind) on trials in which no external pressure to do so was given.
Figure 6.4: Decodes of covert processes. All panels show decode over time. Forced-left trials are displayed in faint blue, forced-right trials in faint red, and the trial of interest in bright color (blue for left choice, red for right). Insets show the final maze configuration and movement. (a) ‘Encouraged switch’ trials. These trials began forced, but became choice trials at the time shown by the black dot. In these two cases, the monkey chose the newly-accessible side, behaviorally indicating a change of mind. (b) ‘Free choice’ trials. In these two trials, both targets were accessible throughout the trial, with no barrier changes occurring. Nonetheless, on these trials the monkey appears to have vacillated. Left panels from dataset J1-2, right panels from dataset J2-1.
6.3.1 Challenges for existing models

Unlike most previous paradigms, information in the decision-maze task is frequently delivered all at once. It is therefore somewhat unclear how to apply models such as integration to a bound, race to threshold, or probabilistic coding. While attractor dynamics seem promising, elaborations may be needed to account for spontaneous vacillation. Either requiring the attractors to be very shallow or assuming high levels of decision noise would seem to predict wide plan distributions, when instead the distributions were typically observed to be well-separated (see background traces in Figure 6.4). However, it certainly remains possible that the decision is being largely made upstream of PMd and M1, and that the decision variable is passed through a strong enough nonlinearity to distort these distributions.

6.3.2 Future directions

Many additional analyses of this data are clearly warranted. These fall into at least three rough categories. First, there are numerous additional questions to answer on the basis of decodes, especially regarding the timecourse of events such as changes of mind or initial choice in forced vs. free choice trials. However, decodes provide only a one-dimensional view of the neural data, while the neural state is closer to 10-dimensional.

As a second category of analyses, the full neural trajectory can be considered. This category includes questions such as whether the trajectories of forced trials resemble those of free choices, what the similarities and differences are between the distributions of plan states for forced and free plans, and whether encouraged switches resemble spontaneous vacillation.

Third, one can ask whether details of the decision process are reflected in the fine structure of the neural state. We know from the behavior that barrier changes often induce changes of mind, but that they do not do so every time. If the neural state is predictive of which trials will end in a change of mind and which will not, the nature of this difference may be deeply informative about the dynamics of decision-making. For instance, the decode axis might reflect ‘confidence’ and thus extreme trials should
result in low probability of a change-of-mind; alternatively, such extreme trials may be far from an attractor and thus indicate lability. Alternatively, if the details of the decision variable are not present in the plan state, this may indicate a greater separation of decision-making and motor planning than might have been expected in such a motoric task.

6.4 Methods

6.4.1 Subjects

Animal protocols were approved by the Stanford University Institutional Animal Care and Use Committee. Subjects were two adult male macaque monkeys (*Macaca mulatta*). Details of their surgical history are described in Chapter 3.

6.4.2 Task design and neural recordings

The basic task design is described in Results. The delay distribution and delay between target onset and barrier change were chosen from independent, truncated exponentials ($\tau_{\text{delay}} = 500$ ms, truncated at 1000 ms; $\tau_{\text{change}} = 500$ ms, truncated at 1200 ms). Although it reduced our useful trial count, this independence was required to ensure that the monkey could not use a barrier change as information about when the Go cue would arrive. Trials of interest were found via backsorting. All starting configurations (difficulty levels) were used with equal frequency. The targets were rewarded equally, and in the case of both options being inaccessible throughout the trial, the monkey was not rewarded.

Neural recordings were obtained using two 96-channel Utah electrode arrays (Blackrock Microsystems, Salt Lake City, UT) implanted in M1 and caudal PMd of each monkey, as estimated from local anatomical landmarks and previous single-electrode recordings. Units were sorted using custom software, and only stable single- and multi-unit activity was included. The array recordings yielded strong, well-tuned preparatory-period activity for arm movements.
6.4.3 Neural decode

Dimensionality was reduced to 12-D using GPFA with 20 ms bin widths. This dimensionality was chosen conservatively based both on previous analysis of PMd activity (Yu et al., 2009) and cross-validation of the present data, which indicated that any dimensionality between $\sim$8-14 would retain most structure without increase in noise. Only data from 300 ms before target onset to 100 ms after Go cue was used. Movement epoch data was excluded because it tends to be much larger and thus dominates the resulting space, reducing the quality of delay epoch trajectory estimation.

As noted in Results, the decode axis was independent at each timepoint, with the axis used determined by the time from target onset. It was trained using only data from trials with a delay of at least 400 ms, and new decode axes were computed at each time point until fewer than half the trials had delays that long.
Chapter 7

Conclusions

7.1 Summary of this thesis

Despite decades of recording neural responses in the motor cortex of awake, behaving monkeys, the most basic question of what the neurons encode has largely escaped answer. In this thesis, a somewhat unorthodox approach was taken: instead of trying to correlate responses of individual neurons to external parameters of movement, an attempt was made to examine the internal dynamics of the neural population response. This approach was motivated by a view of the motor system as a controller for movement, and therefore a desire to discover the operating principles of the motor control machine. The focus was on examining how preparatory activity relates to, and differs from, movement activity, and on the overall dynamics of the system.

In Chapters 2 and 3, perhaps the most basic question possible about preparatory activity was asked: how does preparatory activity differ from movement activity such that one does not cause movement and the other does? The key findings were that this control of functional connectivity does not appear likely to be achieved via some set of inhibitory neurons performing as a gate. Instead, in considering the whole population, it appears that motor cortex ‘knows’ which activity patterns are muscle-potent and which are not (that is, are iso-force). It is thus possible for motor cortex to prepare in the iso-force space. While such a mechanism initially appears more complicated than inhibitory gating, it is in many ways much simpler: in particular, it allows for
a purely linear mapping from neural activity to muscle activity, unlike the nonlinear operation of a gating transfer function. Moreover, such an idea generalizes: functional connectivity of connected brain areas could be regulated similarly.

Chapter 4 then showed that the preparatory state seeds the dynamics present during movement, substantially determining the neural trajectory taken. Critically, those movement-time dynamics appear to be dominated by rotational (oscillatory) patterns, demonstrating that the apparently complex and confusing patterns seen in PSTHs of motor and premotor neurons are actually highly organized in the context of the population. This discovery, besides clarifying an apparently basic operating principle of motor cortex, suggests that a systematic search for coherent dynamics may be warranted in other brain areas that have presented confusing responses.

In order to fully understand the dynamics of the system, however, one must be able to perturb the system in addition to observing it. In preparation to perform such perturbations with cell-type specificity and/or patterning, Chapter 5 presented work in bringing optogenetics to primates. In addition to identifying a set of viruses and promoters that function in monkey, data that optogenetic stimulation acts differently from electrical stimulation was presented.

Finally, Chapter 6 presented preliminary results from ongoing analyses of a decision-making dataset, examining neural dynamics on single trials. It had previously been known behaviorally that humans can exhibit changes of mind when presented with changing external evidence (Resulaj et al., 2009). Here, it was demonstrated that monkeys are a good model for changes of mind, since they too alter their behavior when evidence changes. In exploring the neural data with dynamical systems tools, it also became possible to observe the moment-by-moment evolution of the decision-making process. These analyses showed, for the first time, that monkeys sometimes vacillate in their decision-making in the absence of external reasons to change their minds. This combination of paradigm and tools may furnish a platform to examine the dynamics of the neural system in the context of decision-making, where individual trials might be expected to behave uniquely.
7.2 Future work

The single-trial view of neural trajectories will doubtless support a great deal more discovery. This is especially patent in the decision-making data, where single-trial views have long been desired. A number of questions appear particularly amenable to further analysis. First, how does switching of plans work? Must switches pass through the baseline state, or is there some ‘expressway’ between motor plan states? This answer may tell us quite a bit about the dynamics of the system, and inform us about how motor plans are computed. If switches must pass through baseline, that would help argue for a phasic reset model (e.g., Sussillo and Abbott, 2009) in which a consistent initial state must be achieved before further computation may be performed. A straighter pathway taken between plan states might instead argue for something more like attractor dynamics. It will also be interesting to see whether switch trajectories do or do not closely resemble spontaneous vacillations. It would be easy to imagine that an externally imposed switch creates a strong new attractor in state space while vacillation takes a different course. In the same vein, it is unclear whether the neural trajectories of forced and free choices should be the same, either in terms of timecourse or path.

More than anything, however, the decision-making experiment was designed to examine changes of mind. Specifically, we presented the monkeys with changes in the middle of many trials, but these changes only sometimes induced the monkey to alter his plan. Hopefully, a more complete analysis of the data will reveal signatures in the brain of the monkey’s willingness to switch choices or propensity to stay with the current choice. Such a finding could indicate either a role for these motor areas in the decision-making itself, or at least that they reflect subtleties such as the level of commitment to a choice.

There are also a number of potential directions to take the iso-force result. First, as discussed in Chapter 3, it would be of great interest to determine whether the iso-force mechanism is at play in controlling the functional connectivity between other brain areas. Unfortunately, this is also likely to prove somewhat challenging. In the motor system, it was possible to take advantage of the relatively straightforward relationship
between neural activity and the muscles. While the spinal cord of course intervenes, apparently it is not so nonlinear as to preclude finding the muscle-potent directions in cortical state space. In the case of other brain areas, a great deal of recurrence exists. This is likely to so alter the activity in the downstream area as to make finding the output-potent space of the upstream area difficult. It may therefore prove necessary to identify primary input and primary output neurons in the two areas, mitigating the linearity-distorting effects of recurrent activity.

Second, the idea of sliding along iso-output dimensions might also provide part of a mechanism for incorporating urgency signals into decision-making. That is, in some cases there is pressure to make a decision quickly, and thus elapsed time must taken into account (Cisek et al., 2009). In this kind of situation, one solution might be to bring the neural state to a region of more sharply diverging dynamics as time elapses, without changing the projection onto the decision axis. In addition to being iso-force, then, the brain might make use of the iso-decision space.

Third, the iso-force mechanism may be useful in the context of brain-computer interfaces. One major problem in that field has been difficulty in getting a neurally-controlled cursor to stop when it reaches the target; instead, it tends to wander around the desired target (Velliste et al., 2010). Identifying the iso-force space may make it easier to ignore changes in neural output that are not intended to affect the cursor, permitting better stopping performance.

The rotational dynamics result may be useful to the field of brain-computer interfaces as well. Present methods typically rely on Kalman filters (e.g., Kim et al., 2011; Wu et al., 2006), which do not take advantage of neural dynamics, or at best use highly simplified models of neural dynamics (Srinivasan et al., 2006). If the neural data is in fact evolving in spirals, identifying these dynamics may both reduce model mismatch, and make it easier to guess the intended control signal in the presence of spiking noise.

This discovery of rotational dynamics has neuroscience implications as well. It suggests that a far simpler view of neural responses in many areas may await a search for population dynamics. Much of prefrontal cortex produces a menagerie of complex and heterogeneous responses (Schall, 1997), and might therefore present an excellent
target for a search for low-dimensional population dynamics.

Optogenetics can also be applied to better understand any of the results developed here. The role of inhibitory cortical neurons has not yet been identified, and optogenetics could be used to selectively alter their activity and help elucidate their role in the network. To causally test the idea of an iso-force space, one could use patterned stimulation to push the neural state within the iso-force space or out of it, and expect to see much larger evoked movements when pushing out of the iso-force space. One could test the generality of rotational dynamics by stimulating mid-movement while simultaneously recording, and expect to see the altered dynamical seed produce a different, but still rotational, neural trajectory. For decision-making, one could push the neural state from having prepared for one option to a prepared state for the other option, and see whether this can bias the monkey’s choice. If so, that may imply that circuits as close to the output as PMd and M1 actually play a role in deciding between possible movements.

In summary, this thesis has explored how the motor system selects an action, achieves a computationally beneficial preparatory state without causing premature movement, and generates activity patterns that can drive appropriate muscle output. These findings comprise fundamental components of our understanding of the computations performed by the motor system machine, and point to an exciting new set of scientific questions.
Appendix A

Publications

A.1 Journal Articles


This work appears as Chapter 3. I collected the data, designed and performed all analyses, and wrote the manuscript.


This work appears in condensed and somewhat modified form as Chapter 4. I collected much of the data, and was heavily involved in conception and design of the analyses and in editing the manuscript.


This work does not appear in this dissertation. I provided complex behavior training for one of the monkeys.


This work appears in condensed form as Chapter 5. I collected much of the data, designed and implemented the stimulation protocols, designed and implemented the quantitative and video analyses, and edited the manuscript.


This work does not appear in this dissertation. I provided data and wrote spike-sorting software that was used in performing some analyses.


This work appears as Chapter 2. I collected half the data, designed and performed all analyses, and wrote the manuscript.


Small portions of this work appear as part of Chapter 4. I collected much of the data, and was involved in conception and design of the analyses and in editing the manuscript.

### A.2 Refereed conference articles and abstracts


This work became Kaufman et al. (submitted).

This work became Kaufman et al. (submitted).


This work became Churchland et al. (submitted).


This work became Churchland et al. (submitted).


This work was an invited symposium following from Diester et al. (2011).


This work became Churchland et al. (2010).

A.3 Book chapters


I participated in structuring the chapter, and in editing.
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Hatsopoulos, N. G., 2005. Encoding in the motor cortex: was evarts right after all?
focus on “motor cortex neural correlates of output kinematics and kinetics during

Hatsopoulos, N. G., Xu, Q., Amit, Y., 2007. Encoding of movement fragments in the

Heekeren, H. R., Marrett, S., Ungerleider, L. G., 2008. The neural systems that me-
diate human perceptual decision making. Nat Rev Neurosci 9 (6), 467–79.

Intracellular features predicted by extracellular recordings in the hippocampus in

primary motor cortex to transient and continuous loads during posture. Journal of
Neurophysiology 101 (1), 150–63.

Hikosaka, O., Wurtz, R. H., 1983. Effects on eye movements of a gaba agonist and
antagonist injected into monkey superior colliculus. Brain Res 272 (2), 368–72.

Histed, M. H., Bonin, V., Reid, R. C., 2009. Direct activation of sparse, distributed
populations of cortical neurons by electrical microstimulation. Neuron 63 (4), 508–
22.


Hogan, N., Sternad, D., 2007. On rhythmic and discrete movements: reflections, def-

Horwitz, G. D., Newsome, W. T., 1999. Separate signals for target selection and

Horwitz, G. D., Newsome, W. T., 2001. Target selection for saccadic eye movements:
prelude activity in the superior colliculus during a direction-discrimination task. J
Neurophysiol 86 (5), 2543–58.
REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


